

Interference Effects in Quantum Belief Networks

Catarina Moreira

catarina.p.moreira@ist.utl.pt

Andreas Wichert

andreas.wichert@ist.utl.pt

Instituto Superior Técnico, INESC-ID
Av. Professor Cavaco Silva, 2744-016 Porto Salvo, Portugal

The original publication is available at: Applied Soft Computing, Elsevier
<http://www.sciencedirect.com/science/article/pii/S095741741300238>

Abstract

Probabilistic graphical models such as Bayesian Networks are one of the most powerful structures known by the Computer Science community for deriving probabilistic inferences. However, modern cognitive psychology has revealed that human decisions could not follow the rules of classical probability theory, because humans cannot process large amounts of data in order to make judgements. Consequently, the inferences performed are based on limited data coupled with several heuristics, leading to violations of the law of total probability. This means that probabilistic graphical models based on classical probability theory are too limited to fully simulate and explain various aspects of human decision making.

Quantum probability theory was developed in order to accommodate the paradoxical findings that the classical theory could not explain. Recent findings in cognitive psychology revealed that quantum probability can fully describe human decisions in an elegant framework. Their findings suggest that, before taking a decision, human thoughts are seen as superposed waves that can interfere with each other, influencing the final decision.

In this work, we propose a new Bayesian Network based on the psychological findings of cognitive scientists. We made experiments with two very well known Bayesian Networks from the literature. The results obtained revealed that the quantum like Bayesian Network can affect drastically the probabilistic inferences, specially when the levels of uncertainty of the network are very high (no pieces of evidence observed). When the levels of uncertainty are very low, then the proposed quantum like network collapses to its classical counterpart.

This work was supported by national funds through FCT - Fundação para a Ciência e a Tecnologia, under project PEst-OE/EEI/LA0021/2013

1 Introduction

The problem of violations of the axioms of probability go back to the early 60s. [32] published a work that influenced modern psychology by showing that humans violate the laws of probability theory when making decisions under risk. The principle that humans were constantly violating is defined by *The Sure Thing Principle*. It is a concept widely used in game theory and was originally introduced by [64]. This principle is fundamental in Bayesian probability theory and states that if one prefers action A over B under state of the world X , and if one also prefers A over B under the complementary state of the world X , then one should always prefer action A over B even when the state of the world is unspecified.

Cognitive psychologists A. Tversky and D. Kahneman also explored more situations where classical probability theory could not be accommodated in human decisions. In their pioneering work, [69] realised that the beliefs expressed by humans could not follow the rules of Boolean logic or classical probability theory, because humans cannot process large amounts of data in order to make estimations or judgements. Consequently, the inferences performed are based on limited data coupled with several heuristics, leading to a violation on one of the most important laws in Bayesian theory: the law of total probability.

One of the key differences between classical and quantum theories is the way how information is processed. According to classical decision making, a person changes beliefs at each moment in time, but it can only be in one precise state with respect to some judgement. So, at each moment, a person is favouring a specific belief. The process of human inference deterministically either jumps between definite states or stays in a single definite state across time [17]. Most computer science, cognitive and decision systems are modelled according to this single path trajectory principle. Figure 1 illustrates this idea.

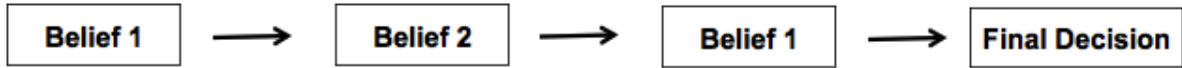


Figure 1: Example of how information is processed in a classical setting. At each time, beliefs can only be in one definite state.

In quantum information processing, on the other hand, information (and consequently beliefs) are modelled via wave functions and therefore they cannot be in definite states. Instead, they are in an indefinite quantum state called the *superposition* state. That is, all beliefs are occurring on the human mind at the same time. According to cognitive scientists, this effect is responsible for making people experience uncertainties, ambiguities or even confusion before making a decision. At each moment, one belief can be more favoured than another, but all beliefs are available at the same time. In this sense, quantum theory enables the modelling of the cognitive system as it was a wave moving across time over a state space until a final decision is made. From this superposed state, uncertainty can produce different waves coming from opposite directions that can crash into each other, causing an interference distribution. This phenomena can never be obtained in a classical setting. Figure 2 exemplifies this. When the final decision is made, then there is no more uncertainty. The wave collapses into a definite state. Thus, quantum information processing deals with both definite and indefinite states [17].

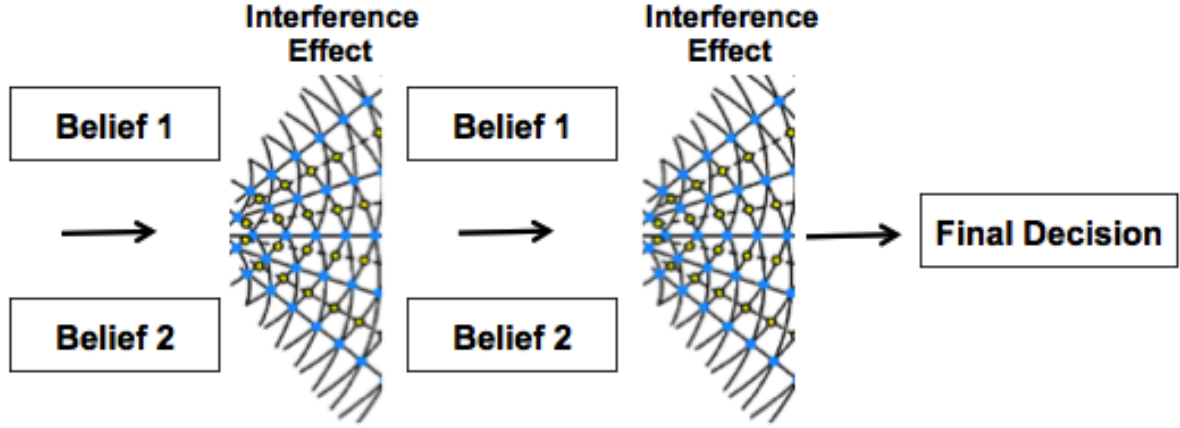


Figure 2: In human decision making, beliefs occur in the human mind at the same time, leading to uncertainty feelings and ambiguity. Beliefs can be represented in superposition states that can generate interferences between them.

1.1 Motivation: Violations in The Two-Stage Gambings

[71] were one of the first researchers to test the veracity of Savage's principle under human cognition in a gambling game. In their experiment, participants were asked at each stage to make the decision of whether or not to play a gamble that has an equal chance of winning \$200 or losing \$100. Figure 3 illustrates the experiment. Three conditions were verified:

1. Participants were informed if they had won the first gamble;
2. Participants were informed if they had lost the first gamble;
3. Participants did not know the outcome of the first gamble;

The two-stage gambling game was one of the first experiments used in order to determine if the Sure Thing Principle would be verified even with people that did not know about the existence of this principle. The results obtained in [71] experiment showed that this principle is constantly being violated and consequently humans do not perform inferences according to the laws of probability theory and Boolean logic.

The overall results revealed that participants who knew that they won the first gamble, decided to play again. Participants who knew that they lost the first gamble, also decided to play again. Through Savage's sure thing principle, it was expected that the participants would choose to play again, even if they did not know the outcome of the first gamble. However, the results obtained revealed something different. If the participants did not know the outcome of the first gamble, then many of them decided not to play the second one.

Several researchers replicated this experiment. The overall results are specified in Table 1.

Why did the findings reported in Table 1 generate so much controversy in the scientific community? Because, the data observed is not in accordance with the classical law of total probability. In Tversky and Shafir's experiment [71], the probability of a participant playing the second gamble, given that the outcome of the first gamble is unknown, $Pr(G|U)$, can be computed through the law of total probability:

$$Pr(G | U) = Pr(W | U) \cdot Pr(G | W) + Pr(L | U) \cdot Pr(G | L) \quad (1)$$

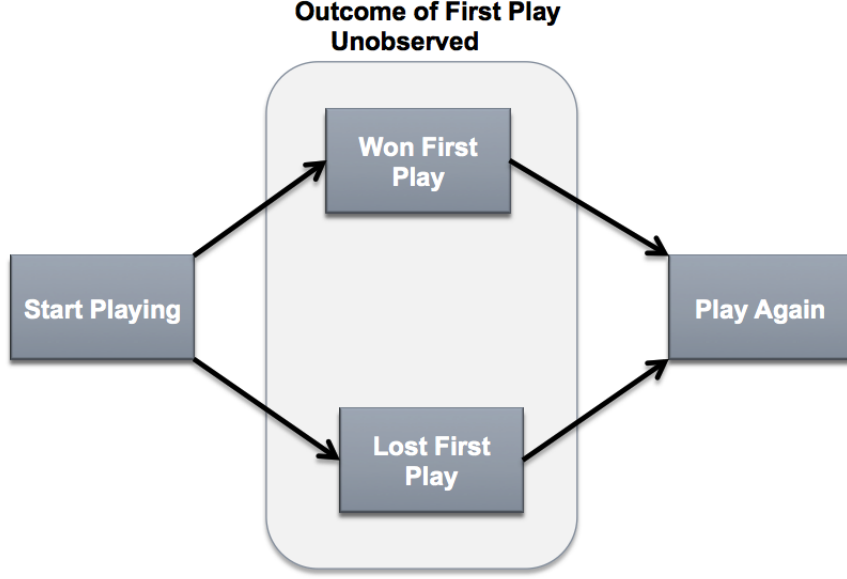


Figure 3: The two-stage gambling experiment proposed by [71]

Literature	Pr(Play Again Won 1st Play)	Pr(Play Again Lost 1st Play)	Pr(Play Again Unknown 1st Play)
[71]	69%	58%	37%
[50]	72%	47%	48%
[51]	63%	45%	41%
Average	68%	50%	42%

Table 1: Observations reported by different works in the literature about the two-step gambling game.

In Equation 1, $Pr(W|U)$ corresponds to the probability of a player winning the first gamble, given that (s)he participated on the game in the first place. $Pr(G|W)$ is the probability of playing the second gamble, given that it is known that the player won the first one. $Pr(L|U)$ corresponds to the probability of losing the first gamble, given that the participant decided to play the game in the first place. And finally, $Pr(G|L)$ is the probability of a participant playing the second gamble, given that it is known that (s)he lost the first one.

Following the law of total probability in Equation 1, the probability of playing the second gamble, given that the player did not know the outcome of the first one, should be between the following values [17]:

$$Pr(G|W) \geq Pr(G|U) \geq Pr(G|L) \quad (2)$$

The findings reported by [71], however, revealed a different relation. Equation 3 demonstrates that this relation is violating one of the most fundamental laws of Bayesian probability theory:

$$Pr(G|W) = 0.69 \geq Pr(G|L) = 0.58 \geq Pr(G|U) = 0.37 \quad (3)$$

[71] explained these findings in the following way: when the participants knew that they won, then they had extra house money to play with and decided to play the second round. If the participants knew that they lost, then they chose to play again with the hope of recovering the

lost money. But, when the participants did not know if they had won or lost the first gamble, then these thoughts, for some reason, did not emerge in their minds and consequently they decided not to play the second gamble. Other works in the literature also replicated this two-stage gambling experiment [65, 50, 51], also reporting similar results to [71]. Their results are summarised in Table 1.

There have been different works in the literature trying to explain and model this phenomena [17, 61, 21]. Although the models in the literature diverge, they all agree in one thing: one cannot use classical probability theory to model this phenomena, since the most important rules are being violated. This two stage gambling game experiment was one of the most important works that motivated the use of different theories outside of classical bayesian theory and boolean logic, more specifically the usage of quantum probability theory.

1.2 Research Questions

Recent findings in the cognitive psychology literature revealed that humans are constantly violating the law of total probability when making decisions under risk [14, 21, 22]. These researchers also showed that quantum probability theory enables the development of decision models that are able to simulate human decisions. Given that most of the systems that are used nowadays are based on Bayesian probability theory, is it possible to achieve better inference mechanisms in these systems using quantum probability theory? For instance, many medical diagnosing systems are based in classical probabilistic graphical models such as Bayesian Networks. Can one achieve better performances in diagnosing patients using quantum probability?

Generally speaking, a Bayesian Network is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph.

There are two main works in the literature that have contributed to the development and understanding of Quantum Bayesian Networks. One belongs to [68] and the other to [53].

In the work of [68], it is argued that any classical Bayesian Network can be extended to a quantum one by replacing real probabilities with quantum complex amplitudes. This means that the factorisation should be performed in the same way as in a classical Bayesian Network. One big problem with Tucci's work is concerned with the inexistence of any methods to set the phase parameters. The author states that, one could have infinite Quantum Bayesian Networks representing the same classical Bayesian Network depending on the values that one chooses to set the parameters. This requires that one knows *a priori* which parameters would lead to the desired solution for each node queried in the network (which we never know).

In the work of [53], the authors argue that, in order to develop a quantum Bayesian Network, it is required a quantum version of probability distributions, quantum marginal probabilities and quantum conditional probabilities. The proposed model fails to provide any advantage relatively to the classical models, because it cannot take into account interference effects between unobserved random variables. In the end, both models provide no advantages in modelling decision making problems that try to predict decisions that violate the laws of total probability.

In this paper, the core of the proposed Bayesian Network is based on the psychological findings uncovered in the works of [21, 22, 17, 61] and on quantum information processing. These authors show that, before taking a decision, human thoughts are seen as superposed waves that can interfere with each other, influencing the final decision. In Bayesian Networks, nodes can either be query variables, evidences or simply unknown. Given that we do not observe the unknown nodes of a Bayesian Network, since we do not know for sure what values

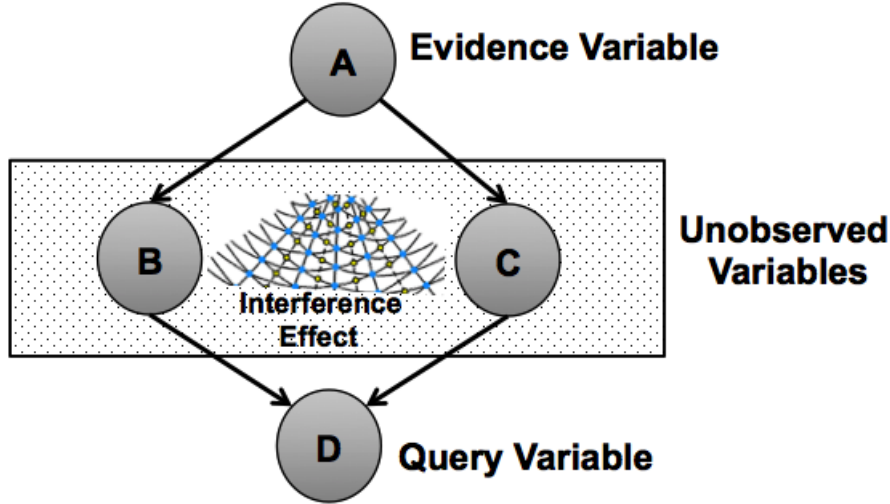


Figure 4: An example of the proposed model. In a quantum Bayesian Network we will assume that if some nodes are unobserved, then the inference process is propagated like a wave through both nodes like a superposed state. Interference effects may arise since the waves can interfere with each other.

they can take, then what would happen to the inference process if these nodes are put in a representation of a quantum superposition and interfere with each other (Figure 4)? Can a better inference process be achieved? These are the main research questions that this paper aims at answering. So far, to the best of our knowledge, there is no previous work in the Computer Science community that attempts to map these psychological findings into computer science decision making systems, such as Bayesian Networks.

In order to validate our hypothesis, we performed experiments with well known classical Bayesian Networks from the literature. We first create a quantum Bayesian Network that can accommodate the paradoxical findings in the two-stage gambling game. We then generalise our quantum Bayesian Network in order to deal with larger and more complex datasets that are used in the literature: the well known Burglar/Alarm Bayesian Network from [63] and the Lung Cancer Bayesian Network from [59].

1.3 Outline

Before describing the proposed model, we first need to introduce some quantum probability concepts for the understanding of this work. Sections 2 and 3 present the main differences between classical and quantum probability theory. Instead of just presenting a set of formulas, we show this difference by means of an illustrative example, just like proposed in [17]. In Section 4, we describe how beliefs can act like waves and interfere with each other. We show mathematically how this interference term can be derived by using well known rules of complex numbers. Section 5 addresses the main works of the literature that contributed for the development of the interference term. It also introduces a new interference formula that will be applied in the proposed quantum probabilistic graphical models. Section 6 presents a comparison between a classical Bayesian Network model against the proposed quantum interference Bayesian Network applied to the problem of two-stage gambles. Section 7 presents another

comparison between the classical and quantum Bayesian Networks, but for a more complex network from the literature. In Section 8, it is made a discussion about the results obtained in the experiments performed in Section 7. Section 9 presents an additional experiment over another Bayesian Network in order to study the impact of the quantum interference parameters in different scenarios. Section 10 presents the most relevant works of the literature. Finally, Section 11 presents the main conclusions of this work.

2 Probability Axioms of Classical and Quantum Theory

In this section, we describe the main differences between classical theory and quantum probability theory through examples. The example analyzed concerns jury duty. Suppose you are a juror and you must decide whether a defendant is guilty or innocent. The following sections describe how the classical and quantum theory evolve in the inference process. All this analysis is based on the book of [17].

2.1 Space

In classical probability theory, events are contained in Sample Spaces. A Sample Space Ω corresponds to the set of all possible outcomes of an experiment or random trial [31]. For example, when judging whether a defendant is guilty or innocent, the sample space is given by $\Omega = \{Guilty, Innocent\}$. Figure 5 presents a diagram showing the sample space of a defendant being guilty or innocent.

In quantum probability theory, events are contained in the so called Hilbert Spaces. A Hilbert Space H can be viewed as a generalisation and extension of the Euclidean space into spaces with any finite or infinite number of dimensions. It can be seen as a vector space of complex numbers and offers the structure of an inner product to enable the measurement of angles and lengths [40]. The space is spanned by a set of orthonormal basis vectors $H = \{Guilty, Innocent\}$. Together, these vectors form a basis for the space. Figure 6 presents a diagram showing the Hilbert space of a defendant being guilty or innocent [17]. Since a Hilbert space enables the usage of complex numbers, then, in order to represent the events *Guilty* and *Innocent*, one would need two dimensions for each event (one for the real part and another for the imaginary part). In quantum theory, one usually ignores the imaginary component in order to be able to visualise geometrically all vectors in a 2-dimensional space.

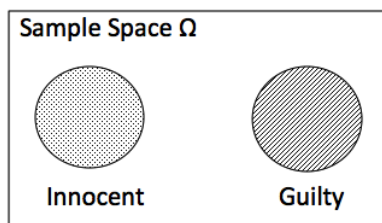


Figure 5: In classical probability theory, results are contained in sample spaces.

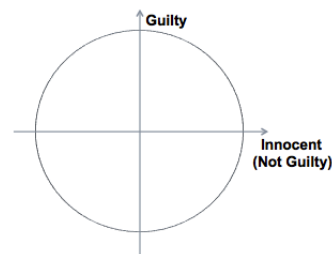


Figure 6: In quantum probability theory, events are spanned by a set of orthonormal basis vectors in a Hilbert Space.

2.2 Events

In classical probability theory, events can be defined by a set of outcomes to which a probability is assigned. They correspond to a subset of the sample space Ω from which they are contained in. Events can be mutually exclusive and they obey to set theory. This means that operations such as intersection or union of events are well defined. Since they respect set theory, the distributive axiom is also defined between sets. In our example, *Guilty* or *Innocent* can be seen as two mutually exclusive events.

According to quantum probability theory, events correspond to a subspace spanned by a subset of the basis vectors contained in the Hilbert Space. Events can be orthogonal, that is, they can be mutually exclusive. Operations such as intersection and union of events are well defined if the events are spanned by the same basis vector [17]. In quantum theory, all events contained in a Hilbert Space are defined through a superposition state which is represented by a state vector S comprising the occurrence of all events. In our example, *Guilty* and *Innocent* correspond to column vectors representing the main axis of the circle in Figure 7. They are defined as follows:

$$Guilty = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad Innocent = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

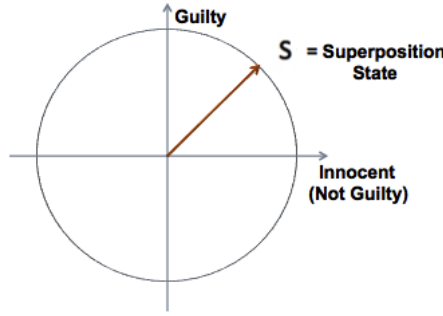


Figure 7: Example of an event represented by a superposition of the states *Guilty* and *Innocent* denoted by the quantum state S .

In Figure 7, the superposition state S can be defined as follows.

$$S = \frac{e^{i\theta_G}}{\sqrt{2}} Guilty + \frac{e^{i\theta_I}}{\sqrt{2}} Innocent \quad (4)$$

In Equation 4, one might be wondering what the $\frac{e^{i\theta}}{\sqrt{2}}$ values mean. They are called *probability amplitudes*. They correspond to the amplitudes of a wave and are described by complex numbers. The $e^{i\theta}$ term is defined as the phase of the amplitude. It can be seen as a shift of the wave. These amplitudes are related to classical probability by taking the squared magnitude of these amplitudes. This is achieved by multiplying the amplitude with its complex conjugate (represented by the symbol $*$).

$$Pr(Guilty) = \left| \frac{e^{i\theta_G}}{\sqrt{2}} \right|^2 = \left(\frac{e^{i\theta_G}}{\sqrt{2}} \right) \cdot \left(\frac{e^{i\theta_G}}{\sqrt{2}} \right)^* = \frac{e^{i\theta_G}}{\sqrt{2}} \cdot \frac{e^{-i\theta_G}}{\sqrt{2}} = e^{i(\theta_G - \theta_G)} \left(\frac{1}{\sqrt{2}} \right)^2 = 0.5 \quad (5)$$

In quantum theory, it is required that the sum of the squared magnitudes of each amplitude equals 1. This axiom is called the normalization axiom and corresponds to the classical theory constraint that the probability of all events in a sample space should sum to one.

$$\left| \frac{e^{i\theta_G}}{\sqrt{2}} \right|^2 + \left| \frac{e^{i\theta_I}}{\sqrt{2}} \right|^2 = 1 \quad (6)$$

2.3 System State

A system state is nothing more than a probability function Pr which maps events into probability numbers, i.e., positive real numbers between 0 and 1.

In classical theory, the system state corresponds to exactly its definition. There is a function that is responsible to assign a probability value to the outcome of an event. If the event corresponds to the sample space, then the system state assigns a probability value of 1 to the event. If the event is empty, then it assigns a probability of 0. In our example, if nothing else is told to the juror, then the probability of the defendant being guilty is $Pr(Guilty) = 0.5$.

In quantum theory, the probability of a defendant being *Guilty* is given by the squared magnitude of the projection from the superposition state S to the subspace containing the observed event *Guilty*. Figure 8 shows an example. If nothing is told to the juror about the guiltiness of a defendant, then according to quantum theory, we start with a superposition state S .

$$S = \frac{e^{i\theta_G}}{\sqrt{2}} Guilty + \frac{e^{i\theta_I}}{\sqrt{2}} Innocent$$

When someone asks whether the defendant is guilty, then we project the superposition state S into the relevant subspace, in this case the *Guilty* subspace (P_G), just like shown in Figure 8. The probability is simply the squared magnitude of the projection, that is:

$$Pr(Guilty) = |P_G|^2 = \left| \frac{e^{i\theta_G}}{\sqrt{2}} \right|^2 = 0.5$$

Which has exactly the same outcome as in the classical theory.

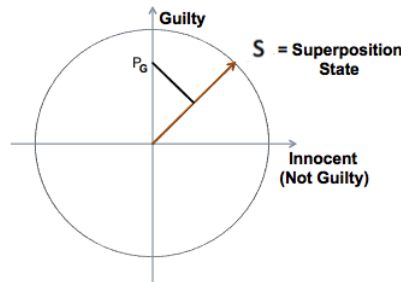


Figure 8: Geometric representation of a projection into the *Guilty* subspace. The computation of the probability starts in the superposition state S . Then, this superposition state is projected into the *Guilty* subspace P_G . The final probability corresponds to the squared magnitude of the projection.

2.4 State Revision

State revision corresponds to the situation where after observing an event, we are interested in observing other events given that the previous one has occurred.

In classical theory, this is addressed through the conditional probability formula $Pr(B|A) = \frac{Pr(A \cap B)}{Pr(A)}$. So, returning to our example, suppose that some evidence has been given to the juror proving that the defendant is actually guilty, then what is the probability of him being innocent? This is computed in the following way.

$$Pr(Innocent|Guilty) = \frac{Pr(Innocent \cap Guilty)}{Pr(Guilty)} = 0$$

Since the events *Guilty* and *Innocent* are mutually exclusive, then their intersection is empty, leading to a zero probability value.

In quantum theory, the state revision is given by first projecting the superposition state S into the subspace representing the observed event. Then, the projection is normalised such that the resulting vector is unit length. Again, if we want to determine the probability of a defendant being innocent, given he was found guilty, the calculations are performed as follows. We first start in the superposition state vector S .

$$S = \frac{e^{i\theta_G}}{\sqrt{2}} \text{Guilty} + \frac{e^{i\theta_I}}{\sqrt{2}} \text{Innocent}$$

Then, we observe that the defendant is guilty, so we project the state vector S into the *Guilty* subspace and normalise the resulting projection.

$$S_G = \frac{(e^{i\theta_G}/\sqrt{2}) \text{Guilty}}{\sqrt{|e^{i\theta_G}/\sqrt{2}|^2}}$$

$$S_G = e^{i\theta_G} \text{Guilty} + 0 \text{Innocent}$$

From the resulting state, we just extract the probability of being innocent by simply squaring the respective probability amplitude. Again, we obtain the same results as the classical theory.

$$Pr(Innocent|Guilty) = |0|^2 = 0$$

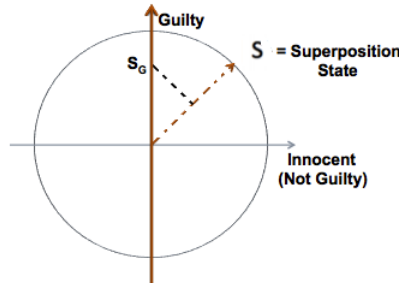


Figure 9: Geometric representation of a state revision. Note that the new state is represented uniquely by the *Guilty* subspace.

3 The Path Trajectory Principle

In order to describe direct dependencies between a set of variables, path diagrams are generally used. This section shows how to compute quantum probabilities in a Markov model using Feynman's path rules, just like presented in the work of [21].

3.1 Single Trajectories

Consider the diagram represented in Figure 10.

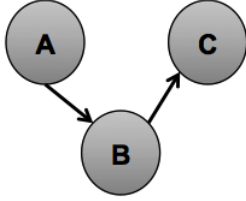


Figure 10: Single path trajectory

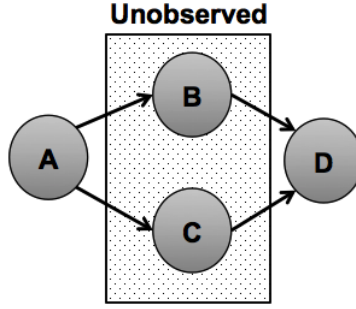


Figure 11: Multiple indistinguishable paths

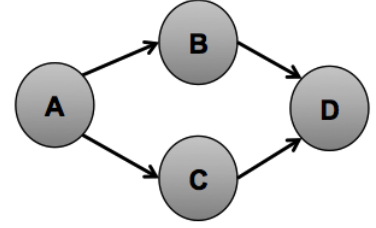


Figure 12: Multiple distinguishable paths

The computation of the probability of transiting from an initial state A to a final state C , transiting from an intermediate state B , can be achieved through a classical Markov model. The probability can be computed by making the product of the individual probabilities for each transition, from one state to another, through the usage of conditional probabilities. According to Figure 10, the probability of transiting from state A , followed by state B and ending in state C , that is, $Pr(A \rightarrow B \rightarrow C)$, is given by:

$$Pr(A \rightarrow B \rightarrow C) = Pr(A) \cdot Pr(B | A) \cdot Pr(C | B) \quad (7)$$

In a quantum path diagram model, the computation of the probabilities for a single path trajectory is similar to the classical Markov model. The calculation can be performed using Feynman's first rule, which asserts that the probability of a single path trajectory consists in the product of the squared magnitudes of the amplitudes for each transition from one state to the next along the path. This means that the quantum probability value of a single path trajectory is the same as the classical Markov probability for the same path. In Equation 8 and through the rest of this work, complex probability amplitudes will be represented by the symbol ψ .

$$Pr(A \rightarrow B \rightarrow C) = |\psi_A|^2 \cdot |\psi_{B|A}|^2 \cdot |\psi_{C|B}|^2 = Pr(A) \cdot Pr(B | A) \cdot Pr(C | B) \quad (8)$$

3.2 Multiple Indistinguishable Trajectories

An indistinguishable path consists in transiting from an initial state A to a final state D by transiting from multiple possible paths without knowing for certain which path was taken to reach the goal state. Figure 11 shows an example of multiple indistinguishable trajectories.

In a classical Markov model, if one does not observe which path was taken to reach the final state D , then one simply computes this probability by summing the individual probabilities of each path. This is known as the single path principle and is in accordance with the law of total probability. So, following Figure 11, in order to reach state D starting in state A , one can take the path $A \rightarrow B \rightarrow D$ or the path $A \rightarrow C \rightarrow D$. The final probability is given by:

$$Pr(A \rightarrow D) = Pr(A) \cdot Pr(B | A) \cdot Pr(D | B) + Pr(A) \cdot Pr(C | A) \cdot Pr(D | C) \quad (9)$$

Quantum probability theory rejects the single path trajectory principle. If one does not observe which path was taken to reach the goal state, then one cannot assume that one of each possible paths was used. Instead, quantum probability argues that, when the path is unobserved, then the goal state can be reached through a superposition of path trajectories. This is known as Feynman's second rule, which states that the amplitude of transiting from an initial state A to a final state D , taking multiple indistinguishable paths, is given by the sum of all amplitudes for each path. This rule is in accordance with the law of total amplitude and the probability is computed by taking the squared magnitude of this sum. This probability is not equal to the classical Markov model.

$$\begin{aligned} Pr(A \rightarrow D) &= |\psi_A \cdot \psi_{B|A} \cdot \psi_{D|B} + \psi_A \cdot \psi_{C|A} \cdot \psi_{D|C}|^2 = \\ &= |\psi_A \cdot \psi_{B|A} \cdot \psi_{D|B}|^2 + |\psi_A \cdot \psi_{C|A} \cdot \psi_{D|C}|^2 + 2 \cdot |\psi_A \cdot \psi_{B|A} \cdot \psi_{D|B}| \cdot |\psi_A \cdot \psi_{C|A} \cdot \psi_{D|C}| \cos(\theta) \end{aligned} \quad (10)$$

The term $\cos(\theta)$ is the inner product between the vectors formed by $|\psi_{B|A} \cdot \psi_{D|B}|$ and $|\psi_{C|A} \cdot \psi_{D|C}|$. It comes from Euler's rule: $\cos(\theta) = (e^{i\theta} + e^{-i\theta})/2$ and corresponds to a quantum interference term that does not exist in classical probability theory. Section 4 details how this term is derived. Since the interference term can lead to estimations with values higher than 1, then it is necessary to normalize this value in order to obtain a probability value.

When the path is observed, quantum probability theory collapses to the classical Markov model. This is known as Feynman's third rule and states that the probability amplitude of observed multiple path trajectories corresponds to the sum of the amplitudes of each individual path. The probabilities are then taken by making the squared magnitude of each individual path. Figure 12 illustrates this example:

$$\begin{aligned} Pr(A \rightarrow D) &= |\psi_A \cdot \psi_{B|A} \cdot \psi_{D|B}|^2 + |\psi_A \cdot \psi_{C|A} \cdot \psi_{D|C}|^2 = \\ &= Pr(A) \cdot Pr(B | A) \cdot Pr(D | B) + Pr(A) \cdot Pr(C | A) \cdot Pr(D | C) \end{aligned} \quad (11)$$

4 The Interference Term

Quantum theory enables the modeling of the decision system as a wave moving across time over a state space until a final decision is made. Under this perspective, interference can be regarded as a chain of waves in a superposition state, coming from different directions. When these waves crash, one can experience a destructive effect (one wave destroys the other) or

a constructive effect (one wave merges with another). In either case, the final probabilities of each wave is affected. Psychological findings showed that this not only occurs in a microscopic scale (such as electrons), but also occurs at a macroscopic setting [44, 61, 22].

In this section we show how to derive the interference term by just taking into account well known properties of complex numbers [48]. The interference term can be derived in two different ways: (1) from the general probability formula of the union of N mutually exclusive events and (2) through the total law of probability.

4.1 Deriving the Interference Term from the Union of N Mutually Exclusive Events

For simplicity, we will start by deriving the interference term for 2 events and then we will generalize for N events. The classical probability formula of the union of two mutual exclusive events is given by:

$$Pr(A \cup B) = Pr(A) + Pr(B) \quad (12)$$

And the relation between a classical probability density function and a quantum probability amplitude is given by Born's rule, that is:

$$Pr(A) = |e^{i\theta_A} \psi_A|^2 \quad (13)$$

Again, $|e^{i\theta_A} \psi_A|^2$ corresponds to the square magnitude of a complex amplitude. It is obtained by multiplying the probability amplitude with its complex conjugate. That is, $|e^{i\theta_A} \psi_A|^2 = e^{i\theta_A} \psi_A e^{-i\theta_A} \psi_A$.

Taking into account Equation 13, one can write a superposition state between two mutual exclusive events A and B in the following way:

$$\psi_{A+B} = e^{i\theta_A} \psi_A + e^{i\theta_B} \psi_B \quad (14)$$

The relation of this superposed state with classical probability theory remains:

$$Pr(A \cup B) = Pr(A) + Pr(B) \propto |e^{i\theta_A} \psi_A + e^{i\theta_B} \psi_B|^2 = |\psi_{A+B}|^2 \quad (15)$$

The quantum counterpart of the classical probability of the union of two mutually exclusive events, when we do not observe them, collapses to Feynmann's second rule, that is:

$$Pr(A \cup B) = |e^{i\theta_A} \psi_A + e^{i\theta_B} \psi_B|^2 = (e^{i\theta_A} \psi_A + e^{i\theta_B} \psi_B) \cdot (e^{i\theta_A} \psi_A + e^{i\theta_B} \psi_B)^*$$

$$|e^{i\theta_A} \psi_A + e^{i\theta_B} \psi_B|^2 = e^{i\theta_A} \psi_A \cdot e^{-i\theta_A} \psi_A + e^{i\theta_A} \psi_A \cdot e^{-i\theta_B} \psi_B + e^{i\theta_B} \psi_B \cdot e^{-i\theta_A} \psi_A + e^{i\theta_B} \psi_B \cdot e^{-i\theta_B} \psi_B$$

$$|e^{i\theta_A} \psi_A + e^{i\theta_B} \psi_B|^2 = |\psi_A|^2 + |\psi_B|^2 + e^{i\theta_A} \psi_A \cdot e^{-i\theta_B} \psi_B + e^{i\theta_B} \psi_B \cdot e^{-i\theta_A} \psi_A \quad (16)$$

The classical probability of the union of two mutual exclusive events is $P(A \cup B) = Pr(A) + Pr(B)$. In Equation 16, the amplitude $|\psi_A|^2$ corresponds to $Pr(A)$ and $|\psi_B|^2$ corresponds to $Pr(B)$. So, what is the additional term, $e^{i\theta_A} \psi_A \cdot e^{-i\theta_B} \psi_B + e^{i\theta_B} \psi_B \cdot e^{-i\theta_A} \psi_A$, that we

derived in Equation 16? This term does not exist in classical probability theory and is called *the interference term*. The interference term can be rewritten like in Equation 17:

$$e^{i\theta_A} \psi_A \cdot e^{-i\theta_B} \psi_B + e^{i\theta_B} \psi_B \cdot e^{-i\theta_A} \psi_A = |\psi_A| |\psi_B| e^{i(\theta_A - \theta_B)} + |\psi_A| |\psi_B| e^{i(\theta_B - \theta_A)} \quad (17)$$

Knowing that,

$$\cos(\theta_1 - \theta_2) = \frac{e^{i(\theta_1 - \theta_2)} + e^{i(-\theta_1 + \theta_2)}}{2}$$

Then Equation 17 can be rewritten as:

$$e^{i\theta_A} \psi_A \cdot e^{-i\theta_B} \psi_B + e^{i\theta_B} \psi_B \cdot e^{-i\theta_A} \psi_A = 2 |\psi_A| |\psi_B| \cos(\theta_A - \theta_B) \quad (18)$$

So, the complete quantum probability formula for the union of 2 mutually exclusive events is given by:

$$\psi_{AB} = |e^{i\theta_A} \psi_A + e^{i\theta_B} \psi_B|^2 = |\psi_A|^2 + |\psi_B|^2 + 2 |\psi_A| |\psi_B| \cos(\theta_A - \theta_B) \quad (19)$$

In the above formula, the angle $\theta_A - \theta_B$ corresponds to the phase of the inner product between $|\psi_A\rangle$ and $|\psi_B\rangle$.

Equation 19 only computes the probability of the union of 2 mutually exclusive events. In classical probability theory, if we want to compute the probability of the union of N mutually exclusive events, then we use a generalization of Equation 12, that is:

$$Pr(A_1 \cup A_2 \cup \dots \cup A_N) = \sum_{i=1}^N Pr(A_i) \quad (20)$$

In quantum theory, we make an analogous calculation. In order to compute the probability of the union of N mutually exclusive events, then one needs to generalize Equation 14. The outcome is given by the following formula:

$$Pr(A_1 \cup A_2 \cup \dots \cup A_N) = |\psi_{A_1} + \psi_{A_2} + \dots + \psi_{A_N}|^2 = \left| \sum_{i=1}^N \psi_{A_i} \right|^2$$

$$|e^{i\theta_1} \psi_{A_1} + e^{i\theta_2} \psi_{A_2} + \dots + e^{i\theta_N} \psi_{A_N}|^2 = \sum_{i=1}^N |\psi_{A_i}|^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N |\psi_{A_i}| |\psi_{A_j}| \cos(\theta_i - \theta_j) \quad (21)$$

4.2 Deriving the Interference Term from the Law of Total Probability

This interference term can also be derived directly from the law of total probability. Suppose that events A_1, A_2, \dots, A_N form a set of mutually disjoint events, such that their union is all in the sample space, Ω , for any other event B . Then, the classical law of total probability can be formulated like in Equation 22.

$$Pr(B) = \sum_{i=1}^N Pr(A_i) Pr(B|A_i) \quad \text{where:} \quad \sum_{i=1}^N Pr(A_i) = 1 \quad (22)$$

The quantum interference law of total probability can be derived through Equation 22 by applying Born's rule (Equation 13). That is:

$$Pr(B) = \left| \sum_{x=1}^N e^{i\theta_x} \psi_{A_x} \psi_{B|A_x} \right|^2 \quad \text{where:} \quad \sum_{x=1}^N \left| e^{i\theta_x} \psi_{A_x} \right|^2 = 1 \quad (23)$$

For simplicity, we will expand Equation 23 for $N = 3$ and only later we will find the general formula for N events:

$$Pr(B) = \left| e^{i\theta_1} \psi_{A_1} \psi_{B|A_1} + e^{i\theta_2} \psi_{A_2} \psi_{B|A_2} + e^{i\theta_3} \psi_{A_3} \psi_{B|A_3} \right|^2 \quad (24)$$

$$Pr(B) = \left(e^{i\theta_1} \psi_{A_1} \psi_{B|A_1} + e^{i\theta_2} \psi_{A_2} \psi_{B|A_2} + e^{i\theta_3} \psi_{A_3} \psi_{B|A_3} \right) \left(e^{i\theta_1} \psi_{A_1} \psi_{B|A_1} + e^{i\theta_2} \psi_{A_2} \psi_{B|A_2} + e^{i\theta_3} \psi_{A_3} \psi_{B|A_3} \right)^* \quad (25)$$

$$Pr(B) = \left(e^{i\theta_1} \psi_{A_1} \psi_{B|A_1} + e^{i\theta_2} \psi_{A_2} \psi_{B|A_2} + e^{i\theta_3} \psi_{A_3} \psi_{B|A_3} \right) \left(e^{-i\theta_1} \psi_{A_1} \psi_{B|A_1} + e^{-i\theta_2} \psi_{A_2} \psi_{B|A_2} + e^{-i\theta_3} \psi_{A_3} \psi_{B|A_3} \right) \quad (26)$$

$$\begin{aligned} Pr(B) = & e^{i\theta_1} \psi_{A_1} \psi_{B|A_1} e^{-i\theta_1} \psi_{A_1} \psi_{B|A_1} + e^{i\theta_1} \psi_{A_1} \psi_{B|A_1} e^{-i\theta_2} \psi_{A_2} \psi_{B|A_2} + e^{i\theta_1} \psi_{A_1} \psi_{B|A_1} e^{-i\theta_3} \psi_{A_3} \psi_{B|A_3} + \\ & + e^{i\theta_2} \psi_{A_2} \psi_{B|A_2} e^{-i\theta_1} \psi_{A_1} \psi_{B|A_1} + e^{i\theta_2} \psi_{A_2} \psi_{B|A_2} e^{-i\theta_2} \psi_{A_2} \psi_{B|A_2} + e^{i\theta_2} \psi_{A_2} \psi_{B|A_2} e^{-i\theta_3} \psi_{A_3} \psi_{B|A_3} + \\ & + e^{i\theta_3} \psi_{A_3} \psi_{B|A_3} e^{-i\theta_1} \psi_{A_1} \psi_{B|A_1} + e^{i\theta_3} \psi_{A_3} \psi_{B|A_3} e^{-i\theta_2} \psi_{A_2} \psi_{B|A_2} + e^{i\theta_3} \psi_{A_3} \psi_{B|A_3} e^{-i\theta_3} \psi_{A_3} \psi_{B|A_3} \end{aligned} \quad (27)$$

Simplifying Equation 27, we obtain:

$$Pr(B) = |\psi_{A_1} \psi_{B|A_1}|^2 + |\psi_{A_2} \psi_{B|A_2}|^2 + |\psi_{A_3} \psi_{B|A_3}|^2 + \text{Interference} \quad (28)$$

In Equation 28, one can see that it is composed by the classical law of total probability and by an interference term. This interference term comes from Equation 27 and corresponds to:

$$\begin{aligned} \text{Interference} = & e^{i\theta_1 - i\theta_2} \psi_{A_1} \psi_{B|A_1} \psi_{A_2} \psi_{B|A_2} + e^{i\theta_2} \psi_{A_2} \psi_{B|A_2} e^{-i\theta_1} \psi_{A_1} \psi_{B|A_1} + \\ & + e^{i\theta_1 - i\theta_3} \psi_{A_1} \psi_{B|A_1} \psi_{A_3} \psi_{B|A_3} + e^{i\theta_3 - i\theta_1} \psi_{A_3} \psi_{B|A_3} \psi_{A_1} \psi_{B|A_1} + \\ & + e^{i\theta_2 - i\theta_3} \psi_{A_2} \psi_{B|A_2} \psi_{A_3} \psi_{B|A_3} + e^{i\theta_3 - i\theta_2} \psi_{A_3} \psi_{B|A_3} \psi_{A_2} \psi_{B|A_2} \end{aligned} \quad (29)$$

Knowing that

$$\cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2} \Rightarrow \cos(\theta_1 - \theta_2) = \frac{e^{i\theta_1 - i\theta_2} + e^{i\theta_2 - i\theta_1}}{2}$$

Then, Equation 28 becomes

$$Pr(B) = |\psi_{A_1} \psi_{B|A_1}|^2 + |\psi_{A_2} \psi_{B|A_2}|^2 + |\psi_{A_3} \psi_{B|A_3}|^2 + 2\psi_{A_1} \psi_{B|A_1} \psi_{A_2} \psi_{B|A_2} \cos(\theta_1 - \theta_2) + \\ + 2\psi_{A_1} \psi_{B|A_1} \psi_{A_3} \psi_{B|A_3} \cos(\theta_1 - \theta_3) + 2\psi_{A_2} \psi_{B|A_2} \psi_{A_3} \psi_{B|A_3} \cos(\theta_2 - \theta_3) \quad (30)$$

Generalizing Equation 30 for N events, the final probabilistic interference formula, derived from the law of total probability, is given by:

$$Pr(B) = \sum_{i=1}^N |\psi_{A_i} \psi_{B|A_i}|^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \psi_{A_i} \psi_{B|A_i} \psi_{A_j} \psi_{B|A_j} \cos(\theta_i - \theta_j) \quad (31)$$

Following Equation 31, when $\cos(\theta_i - \theta_j)$ equals zero, then it is straightforward that quantum probability theory converges to its classical counterpart, because the interference term will be zero.

For non-zero values, Equation 31 will produce interference effects that can affect destructively the classical probability (when interference term is smaller than zero) or constructively (when it is bigger than zero). Additionally, Equation 31 will lead to a large amount of θ parameters when the number of events increases. For N binary random variables, we will end up with 2^N parameters to tune.

5 The Role of the Interference Term in the Literature

In the 20th century, the physicist Max Born proposed a problem related to quantum probabilities, which was later known as the *The Inverse Born Problem*. The problem consisted in constructing a probabilistic representation of data from different sources (physics, psychology, economy, etc), by a complex probability amplitude, which could match Born's rule (already presented in Equation 13).

The probabilistic interference formula for the law of total probability, derived in the previous section (Equation 31), can be seen as an answer to the Inverse Born Problem. The most important works in the literature that contributed for the derivation of this interference term, through the law of total probability, correspond to the works of A. Khrennikov [47, 44, 48, 39, 43, 24]. These authors address the interference term as $|\lambda(B|A)|$. In the situations where $|\lambda(B|A)| \leq 1$, then one can apply the trigonometric formula derived in Equation 31. However, there are some data where this condition is not verified. Therefore, when $|\lambda(B|A)| \geq 1$, the authors propose the usage of a hyperbolic interference term, to act like an upper boundary in order to constraint the probability value to a maximum value of 1. This would require the usage of Hyperbolic Hilbert Spaces instead of the complex ones.

In this paper, we argue that there is no need to represent probabilities in a Hyperbolic Hilbert Space in order to avoid non-probability values. Since we will be dealing with probabilistic graphical models, we will always be required to normalise the probability amplitudes when performing probabilistic inferences. For this work, we follow the same probabilistic paradigm used in traditional Bayesian Networks. Thus, we constrain Equation 32 and Equation 33 by a normalisation factor α that will guarantee that the computed values will always be probabilities lesser or equal than one. This normalisation factor corresponds to Feynman's conjecture [33] that an electron can follow any path. Thus, in order to compute the probability $Pr(B)$ that a

particle ends up at a point B , one must sum over all possible paths that the particle can go through. Since the interference term can lead to estimations with values higher than 1, then it is necessary to normalise in order to obtain a probability value.

$$Pr(B) = \alpha \left[\sum_{i=1}^N |\psi_{A_i} \psi_{B|A_i}|^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \psi_{A_i} \psi_{B|A_i} \psi_{A_j} \psi_{B|A_j} \cos(\theta_i - \theta_j) \right] \quad \text{where} \quad \alpha = \frac{1}{Pr(B) + Pr(\neg B)} \quad (32)$$

$$Pr(A_1 + \dots + A_N) = \alpha \left[\sum_{i=1}^N |\psi_{A_i}|^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N |\psi_{A_i}| |\psi_{A_j}| \cos(\theta_i - \theta_j) \right] \quad (33)$$

$$\text{where } \alpha = \frac{1}{Pr(A_1 + \dots + A_N) + Pr(\neg A_1 + \dots + \neg A_N)}$$

Through the triangular inequality, we can also prove that the proposed interference term is also always positive. This way, the axioms of probability theory that state that $0 \leq Pr(A) \leq 1$ will always be satisfied. By applying the triangular inequality, one can easily demonstrate that Equation 33 and, consequently, Equation 32 have always to be positive.

Through the triangular inequality, Equation 33 can be related to:

$$|\psi_{A_1} + \psi_{A_2} + \dots + \psi_{A_N}|^2 \leq (|\psi_{A_1}| + |\psi_{A_2}| + \dots + |\psi_{A_N}|)^2 \quad (34)$$

$$\sum_{i=1}^N |\psi_{A_i}|^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N |\psi_{A_i}| |\psi_{A_j}| \cos(\theta_i - \theta_j) \leq \sum_{i=1}^N |\psi_{A_i}|^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N |\psi_{A_i}| |\psi_{A_j}| \quad (35)$$

$$\sum_{i=1}^{N-1} \sum_{j=i+1}^N |\psi_{A_i}| |\psi_{A_j}| \cos(\theta_i - \theta_j) \leq \sum_{i=1}^{N-1} \sum_{j=i+1}^N |\psi_{A_i}| |\psi_{A_j}| \quad (36)$$

$$- \sum_{i=1}^{N-1} \sum_{j=i+1}^N |\psi_{A_i}| |\psi_{A_j}| \cos(\theta_i - \theta_j) + \sum_{i=1}^{N-1} \sum_{j=i+1}^N |\psi_{A_i}| |\psi_{A_j}| \geq 0 \quad (37)$$

$$\sum_{i=1}^{N-1} \sum_{j=i+1}^N |\psi_{A_i}| |\psi_{A_j}| [1 - \cos(\theta_i - \theta_j)] \geq 0 \quad (38)$$

The maximum value that $\cos(\theta_i - \theta_j)$ can have is 1 and the minimum value is -1 . So, the minimum value that the term $1 - \cos(\theta_i - \theta_j)$ can have is 0 (when the cosine achieves its maximum value). Given that $\psi_{A_1}, \psi_{A_2}, \dots, \psi_{A_N}$ are always positive numbers, then it is straightforward that Equation 32 has to be always positive.

6 Classical vs Quantum Bayesian Network to Model Two Stage Gambles

6.1 Classical Bayesian Networks

A classical Bayesian Network can be defined by a directed acyclic graph structure in which each node represents a different random variable from a specific domain and each edge represents

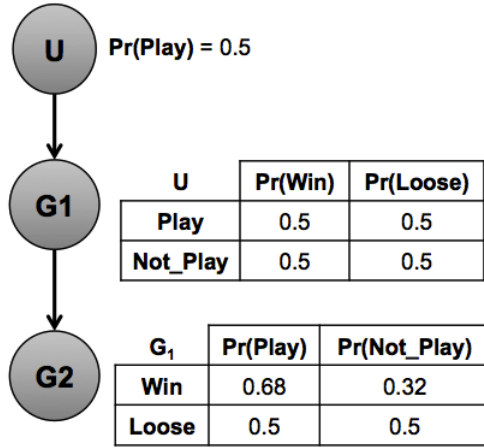


Figure 13: Classical Bayesian Network for the Two-Stage Gambling Game

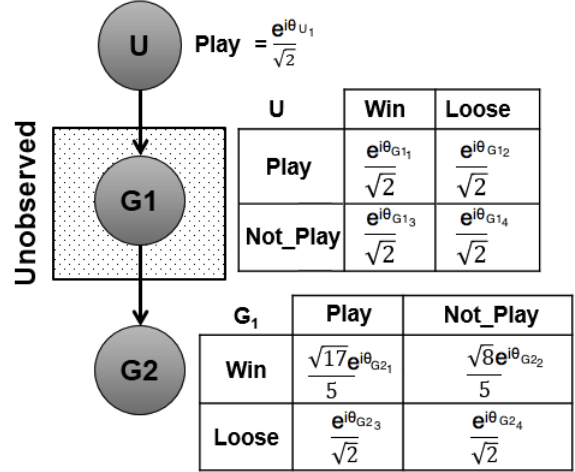


Figure 14: Quantum Bayesian Network for the Two-Stage Gambling Game

a direct influence from the source node to the target node. The graph represents independence relationships between variables and each node is associated with a conditional probability table which specifies a distribution over the values of a node given each possible joint assignment of values of its parents. This idea of a node depending directly from its parent nodes is the core of Bayesian Networks. Once the values of the parents are known, no information relating directly or indirectly to its parents or other ancestors can influence the beliefs about it [49].

6.1.1 Classical Conditional Independence

Associated to Bayesian Networks there is always the concept of conditional independence. Two random variables X and Y are conditionally independent given a third random variable Z if and only if they are independent in their conditional probability distribution given Z . In other words, X and Y are conditionally independent given Z , $(X = x \perp Y = y|Z)$, if and only if, given any value of Z , the probability distribution of X is the same for all values of Y and the probability distribution of Y is the same for all values of X .

This means that an independence statement over random variables is a universal quantification over all possible values of random variables [49]. Therefore, a probability distribution Pr satisfies $(X \perp Y|Z)$ if and only if:

$$Pr(X, Y|Z) = Pr(X|Z)Pr(Y|Z) \quad (39)$$

6.1.2 Classical Random Variables

In classical probability theory, a random variable X is defined by a function that associates a value to each outcome in the sample space Ω , $X : \Omega \rightarrow \mathbb{R}$.

6.1.3 Example of Application in the Two-Stage Gambling Game

In the two-stage gambling game, the random variables correspond to the nodes and their respective conditional probability tables of the Bayesian Network in Figure 13. That is, the variable

U corresponds to a player willing or not to participate in the game. Variable G_1 corresponds to the variable winning or losing the first gamble. Variable G_2 corresponds to a player playing or not the second gamble.

6.1.4 Classical Full Joint Distributions

In classical probability theory, the full joint distribution over a set of n random variables $\mathcal{X} = \{X_1, X_2, \dots, X_n\}$ defined over the same sample space, $Pr(X_1, X_2, \dots, X_n)$, is the distribution that assigns probabilities to events that are specified in terms of these random variable [49]. Then, the full joint distribution of a Bayesian Network, where X is the list of variables, is given by [63]:

$$Pr(X_1, \dots, X_n) = \prod_{i=1}^n Pr(X_i | Parents(X_i)) \quad (40)$$

6.1.5 Example of Application in the Two-Stage Gambling Game

Using Equation 40, the full joint distribution of Figure 13 corresponds to the calculations discriminated in Tables 2.

U	G ₁	G ₂	Pr(U, G ₁ , G ₂)
Play	Win	Play	$0.5 \times 0.5 \times 0.68 = 0.17$
Play	Win	Not_Play	$0.5 \times 0.5 \times 0.32 = 0.08$
Play	Lose	Play	$0.5 \times 0.5 \times 0.5 = 0.125$
Play	Lose	Not_Play	$0.5 \times 0.5 \times 0.5 = 0.125$
Not_Play	Win	Play	$0.5 \times 0.5 \times 0.68 = 0.17$
Not_Play	Win	Play	$0.5 \times 0.5 \times 0.32 = 0.08$
Not_Play	Lose	Play	$0.5 \times 0.5 \times 0.5 = 0.125$
Not_Play	Lose	Play	$0.5 \times 0.5 \times 0.5 = 0.125$

Table 2: Full joint distribution of the Bayesian Network in Figure 13

6.1.6 Classical Marginalization

Given a query random variable X and let Y be the unobserved variables in the network, the marginal distribution of X is simply the probability distribution of X averaging over the information about Y . The marginal probability for discrete random variables, can be defined by Equation 41. The summation is over all possible y , i.e., all possible combinations of values of the unobserved variables y . The term α corresponds to a normalization factor for the distribution $Pr(X)$ [63].

$$Pr(X = x) = \alpha \sum_y Pr(X = x, Y = y) = \alpha \sum_y Pr(X = x | Y = y) Pr(Y = y), \text{ where } \alpha = \frac{1}{\sum_{x \in X} Pr(X = x)} \quad (41)$$

6.1.7 Example of Application in the Two-Stage Gambling Game

After computing the full joint distribution, we need to sum out all the variables that are unknown, in this case, the variable corresponding to the outcome of the first gamble: G_1 . This is achieved by applying the marginal probability formula in Equation 41.

$$Pr(G_2 = Play \mid U = Play) = \alpha \sum_{g \in G_1} Pr(U = Play, G_1 = g, G_2 = Play)$$

$$Pr(G_2 = Play \mid U = Play) = \alpha Pr(U = Play) \sum_{g \in G_1} Pr(G_1 = g \mid U = Play) Pr(G_2 = Play \mid G_1 = g)$$

$$Pr(G_2 = Play \mid U = Play) = \alpha Pr(U = Play) [Pr(G_1 = win \mid U = Play) Pr(G_2 = Play \mid G_1 = win) + Pr(G_1 = lose \mid U = Play) Pr(G_2 = Play \mid G_1 = lose)]$$

$$Pr(G_2 = Play \mid U = Play) = \alpha 0.295$$

$$Pr(G_2 = Not_Play \mid U = Play) = \alpha 0.205$$

The parameter α corresponds to a normalisation factor and, for this example, is given by:

$$\alpha = \frac{1}{Pr(G_2 = Play \mid U = Play) + Pr(G_2 = Not_Play \mid U = Play)} = \frac{1}{0.295 + 0.205} = \frac{1}{0.5}$$

So, the final normalised classical probabilities for the two-stage gambling game correspond to:

$$Pr(G_2 = play \mid U = play) = 0.59$$

$$Pr(G_2 = Not_Play \mid U = Play) = 0.41$$

The probabilities computed are not in accordance with the probabilistic findings reported by [71], because it was empirically observed that $Pr(G_2 = play \mid U = play) = 0.42$. Therefore, a classical Bayesian Network can never be used to model such experiments, because of the relation already presented in Equation 3.

6.2 The Quantum Interference Bayesian Network

Following the work of [53], a Quantum Bayesian Network can be defined by a pair (G, ρ_v) , where $G = (V, E)$ is a directed acyclic graph, and each vertex $v \in V$ is associated with a quantum system with a Hilbert space H_v and ρ_v is a quantum state on $H_v = H_{v1} \otimes H_{v2} \otimes \dots \otimes H_{vn}$. The state ρ_v satisfies the same conditional independence constraints as in a classical Bayesian Network. Note that the definition of a classical Bayesian Network can be directly obtained by replacing the word quantum system by random variable.

The symbol \otimes is defined by *tensor product* and corresponds to a mathematical method that enables the construction of a Hilbert space from the combination of individual Hilbert spaces. Suppose that we have 2 different 2-dimensional Hilbert spaces H_x and H_y , where $H_x = \{x_1, x_2\}$ and $H_y = \{y_1, y_2\}$. Then, their tensor product would be:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \otimes \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 y_1 \\ x_1 y_2 \\ x_2 y_1 \\ x_2 y_2 \end{bmatrix}$$

6.2.1 Quantum Random Variables

In quantum theory, random variables are associated to a set of N quantum systems $V = \{v_1, v_2, \dots, v_N\}$, each associated with a Hilbert space with a specific dimension[53]. Consequently, all values contained in the conditional probability tables associated to the random variables are complex numbers.

6.2.2 Example of Application in the Two-Stage Gambling Game

Each node on the Bayesian Network in Figure 14 can be seen as a subsystem belonging to a specific Hilbert space. For instance, the first node U can be represented in a Hilbert subspace H_u in the following way:

$$U = e^{i\theta_U} \frac{1}{\sqrt{2}} \text{Play} + e^{i\theta_U} \frac{1}{\sqrt{2}} \text{Not_Play}$$

Where Play and Not_Play are column vectors corresponding to the basis of the subspace H_u :

$$\text{Play} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{Not_Play} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Since our goal is to compare the quantum Bayesian Network with its classical counterpart, we will convert the conditional probability tables in Figure 13 into conditional amplitude tables. That is, we simply convert classical probabilities into complex amplitudes through the relation in Equation 13.

$$Pr(A) = |e^{i\theta_A} \psi_A|^2 \rightarrow \psi_A = e^{i\theta_A} \sqrt{Pr(A)} \quad (42)$$

6.2.3 Quantum State

The representation of a general state in a Bayesian Network can be described by a bipartite state. Suppose that $H = H_X \otimes H_Y \otimes H_Z$ is a Hilbert space defined by the composition of three Hilbert spaces H_X , H_Y and H_Z . Then, a quantum state S_{XYZ} is designated bipartite if it can be specified with respect to the random variables X , Y and Z . For X_i , Y_j and Z_k as basis in H_X , H_Y and H_Z , respectively, the bipartite state is given by Equation 43.

$$S_{XYZ} = \sum_{i,j,k} \psi_{X_i} \psi_{Y_j} \psi_{Z_k} X_i \otimes Y_j \otimes Z_k \quad (43)$$

In Equation 43, $\psi_{X_i} \psi_{Y_j} \psi_{Z_k}$ corresponds to the amplitudes of states X_i , Y_j and Z_k , respectively. The states X_i , Y_j and Z_k correspond to column vectors representing basis vectors:

$$X_0 = Y_0 = Z_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad X_1 = Y_1 = Z_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

6.2.4 Example of Application in the Two-Stage Gambling Game

The general quantum state represented by the Bayesian Network in Figure 14 is given by quantum bipartite states. Reformulating Equation 43 for the problem of two step gambles, we obtain:

$$S_{U,G1,G2} = \sum_{ijk} \psi_{U_i} \psi_{G1_j} \psi_{G2_k} U_i \otimes G1_j \otimes G2_k \quad (44)$$

For simplicity, we will write $\psi_{Ui}\psi_{G1j}\psi_{G2k}$ as ψ_{ijk} . So, expanding Equation 44,

$$\begin{aligned}
S_{U,G1,G2} = & \psi_{000} U_0 G1_0 G2_0 e^{i(\theta_{U0}+\theta_{G10}+\theta_{G20})} + \psi_{001} U_0 G1_0 G2_1 e^{i(\theta_{U0}+\theta_{G10}+\theta_{G21})} + \psi_{010} U_0 G1_1 G2_0 e^{i(\theta_{U0}+\theta_{G11}+\theta_{G20})} \\
& + \psi_{011} U_0 G1_1 G2_1 e^{i(\theta_{U0}+\theta_{G11}+\theta_{G21})} + \psi_{100} U_1 G1_0 G2_0 e^{i(\theta_{U1}+\theta_{G10}+\theta_{G20})} + \psi_{101} U_1 G1_0 G2_1 e^{i(\theta_{U1}+\theta_{G10}+\theta_{G21})} + \\
& + \psi_{110} U_1 G1_1 G2_0 e^{i(\theta_{U1}+\theta_{G11}+\theta_{G20})} + \psi_{111} U_1 G1_1 G2_1 e^{i(\theta_{U1}+\theta_{G11}+\theta_{G21})} \quad (45)
\end{aligned}$$

Note that $U_i G1_j G2_k$ are basis column vectors representing the axis of the Hilbert Space,

$$U_0 G1_0 G2_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, U_0 G1_0 G2_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, U_0 G1_1 G2_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \dots U_1 G1_1 G2_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

Following the quantum Bayesian Network in Figure 14, we compute the values of $\psi_{Ui}\psi_{G1j}\psi_{G2k}$ by multiplying the correspondent values in the conditional probability tables. For example, ψ_{000} corresponds to the product of the variables $U = \text{Play}$ with $G1 = \text{Win} \mid U = \text{Play}$ with $G2 = \text{Play} \mid G1 = \text{Win}$. On the other hand, the entry ψ_{110} corresponds to the product of the variables $U = \text{Not Play}$ with $G1 = \text{Lose} \mid U = \text{Not Play}$ with $G2 = \text{Play} \mid G1 = \text{Lose}$. Replacing Equation 45 by the conditional probability values in Figure 14, we obtain:

$$\begin{aligned}
S_{U,G1,G2} = & e^{i(\theta_{U0}+\theta_{G10}+\theta_{G20})} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \frac{\sqrt{17}}{5} U_0 G1_0 G2_0 + e^{i(\theta_{U0}+\theta_{G10}+\theta_{G21})} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \frac{\sqrt{8}}{5} U_0 G1_0 G2_1 + e^{i(\theta_{U0}+\theta_{G11}+\theta_{G20})} \\
& + e^{i(\theta_{U0}+\theta_{G11}+\theta_{G21})} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} U_0 G1_1 G2_1 + e^{i(\theta_{U1}+\theta_{G10}+\theta_{G20})} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \frac{\sqrt{17}}{5} U_1 G1_0 G2_0 + e^{i(\theta_{U1}+\theta_{G10}+\theta_{G21})} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \frac{\sqrt{8}}{5} U_1 G1_0 G2_1 \\
& + e^{i(\theta_{U1}+\theta_{G11}+\theta_{G20})} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} U_1 G1_1 G2_0 + e^{i(\theta_{U1}+\theta_{G11}+\theta_{G21})} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} U_1 G1_1 G2_1
\end{aligned}$$

$$\begin{aligned}
S_{U,G1,G2} = & 0.4123 U_0 G1_0 G2_0 + 0.2828 U_0 G1_0 G2_1 + 0.3536 U_0 G1_1 G2_0 + 0.3536 U_0 G1_1 G2_1 + \\
& + 0.4123 U_1 G1_0 G2_0 + 0.2828 \frac{\sqrt{8}}{5} U_1 G1_0 G2_1 + 0.3536 U_1 G1_1 G2_0 + 0.3536 U_1 G1_1 G2_1
\end{aligned}$$

Note that the sum of the squares of all probability amplitudes $\psi_{Ui}\psi_{G1j}\psi_{G2k}$ sum to 1,

$$\sum_{ijk} |\psi_{Ui}\psi_{G1j}\psi_{G2k}|^2 = 1$$

This bipartite state represents a quantum superposition over all possible states. We can think of this as various wave functions that are occurring at the same time.

6.2.5 Quantum Full Joint Distribution

In quantum probability theory, a full joint distribution is given by a *density matrix*. This matrix provides the probability distribution of all states that a Bayesian Network can have. In our quantum Bayesian Network model, the density matrix ρ corresponds to the multiplication of the bipartite state described in Equation 43 with itself (the symbol \dagger corresponds to the conjugate transpose):

$$\rho = S_{XYZ} S_{XYZ}^\dagger \quad (46)$$

The reason why one multiplies the same state with itself is to obtain the probability value out of the amplitude. Note that the bipartite state contains amplitudes instead of probability values. Knowing that the probability value is obtained by taking the squared magnitude of the amplitude, then, by multiplying a state with its conjugate transpose, one can obtain the full joint probability distribution.

6.2.6 Example of Application in the Two-Stage Gambling Game

From the bipartite state, one can compute the density matrix by applying Equation 46 to this calculation. The density matrix will be useful for later calculations and enables the calculation of the probability distribution of the bipartite state.

$$\rho_{UG1G2} = S_{UG1G2} \cdot S_{UG1G2}^\dagger \quad (47)$$

In this two step gambling game, Equation 47 produces an 8 x 8 density matrix:

$$\rho_{UG1G2} = \begin{bmatrix} |\psi_{000}|^2 & 0 & 0 & \dots & 0 \\ 0 & |\psi_{001}|^2 & 0 & \dots & 0 \\ 0 & 0 & |\psi_{010}|^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & |\psi_{111}|^2 \end{bmatrix} = \begin{bmatrix} 0.1700 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0800 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1250 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1250 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.1700 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0800 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1250 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1250 \end{bmatrix} \quad (48)$$

6.2.7 Quantum Marginalization

The concept of quantum marginalization is analogous to the one in classical probability theory. Given two quantum random variables X and Y , the general idea is to compute the average of the probability distribution of X over the information about Y . This is performed by using the *partial trace* operator, which basically consists in accessing certain positions of the density matrix ρ .

$$X_{ij} = \alpha \sum_{y \in Y} \rho[iy, jy] \quad (49)$$

In Equation 49, the parameter α corresponds to the normalization factor, which is also present in the classical Bayesian Network inference.

6.2.8 Example of Application in the Two-Stage Gambling Game

Considering the two-stage gambling game that we are analyzing, imagining that we want to determine the probability of a participant playing the second gamble, $Pr(G2_{00})$, given that we verified that the player participated in the game in the first place ($u = 0$). That is, we will be summing out variable $G1$. Through Equation 49, one would proceed in the following way:

$$G2_{ij} = G2_{00} = \alpha \sum_{g \in G1} \rho[ugi, ugj] = \alpha(\rho[000, 000] + \rho[010, 010]) = \alpha 0.2950 \rightarrow 0.59$$

And in the same way for $G2_{11}$:

$$G2_{11} = \alpha \sum_{g \in G1, u \in U} \rho[ugi, ugj] = \alpha(\rho[001, 001] + \rho[011, 011]) = \alpha 0.2050 \rightarrow 0.41$$

Note that the indexes of the density matrix are encoded. The assignment 000 corresponds to index 0, the assignments 001 to index 1, ..., and the assignment 111 corresponds to the index 8.

The normalised results correspond to the same probabilities obtained in classical theory. Therefore, we need to incorporate the interference terms found in cognitive psychology into the quantum marginalization formula in order to obtain different results.

6.2.9 Quantum Marginalization with Interference

A quantum interference effect will occur when performing the marginalization of a random variable. If we *do not observe* a random variable, then it can be represented by superposition and perform interferences in other random variables, changing the final outcome.

The quantum interference marginalization formula proposed in this work consists in merging the equation that presents the interference term (Equation 21) with the formula of quantum marginalization (Equation 49). This leads to Equation 50:

$$X_{ij} = \alpha \left| \sum_{y \in Y} \rho[iy, jy] \right|^2$$

$$X_{ij} = \alpha \left(\sum_i^N |\psi_i|^2 + 2 \sum_i^{N-1} \sum_{j=i+1}^N |\psi_i| |\psi_j| \cos(\theta_i - \theta_j) \right), \text{ where } |\psi_i|^2 = \sum_{y \in Y} \rho[iy, jy] \quad (50)$$

6.2.10 Example of Application in the Two-Stage Gambling Game

Following the experiment performed by [71], we want to determine the probability of a participant playing the second gamble, given that (s)he does not know the outcome of the first gamble.

Through Equation 50, this can be computed in the following way:

$$G2_{ij} = \alpha \left| \sum_{u \in U, g \in G1} \rho[iug, jug] \right|^2$$

Given that we want to know the probability of a participant playing the second gamble, then we know for certain that (s)he accepted to play. Thus, we will fix the variable U , representing that a participant accepted to play the game in the first place, $U = 0$:

$$G2_{ij} = \alpha \left| \sum_{g \in G1} \rho[i0g, j0g] \right|^2 \quad (51)$$

Expanding Equation 51,

$$G2_{00} = \alpha |\rho[000, 000] + \rho[010, 010]|^2$$

$$G2_{00} = \alpha |0.17 + 0.125|^2 = \alpha(0.17 + 0.125 + 2\sqrt{0.17}\sqrt{0.125}\cos(\theta_1 - \theta_2)) \quad (52)$$

Computing the probability of a participant deciding to not play the second gamble, given that he does not know the outcome of the first play, we obtain:

$$G2_{11} = \alpha |\rho[001, 001] + \rho[011, 011]|^2$$

$$G2_{11} = \alpha |0.08 + 0.125|^2 = \alpha(0.08 + 0.125 + 2\sqrt{0.08}\sqrt{0.125}\cos(\theta_1 - \theta_2)) \quad (53)$$

The normalization factor α is computed by summing the results when $G2 = Play$ and $G2 = Not_Play$, that is, summing Equation 52 with Equation 53:

$$\alpha = \frac{1}{0.5 + 2\sqrt{0.17}\sqrt{0.125}\cos(\theta_1 - \theta_2) + 2\sqrt{0.08}\sqrt{0.125}\cos(\theta_1 - \theta_2)} = \frac{1}{0.5 + 0.4915\cos(\theta_1 - \theta_2)}$$

Then, the normalized results are given by:

$$G2_{00} = \frac{0.2950 + 2\sqrt{0.17}\sqrt{0.125}\cos(\theta_1 - \theta_2)}{0.5 + 0.4915\cos(\theta_1 - \theta_2)} \quad (54)$$

$$G2_{11} = \frac{0.2050 + 2\sqrt{0.08}\sqrt{0.125}\cos(\theta_1 - \theta_2)}{0.5 + 0.4915\cos(\theta_1 - \theta_2)} \quad (55)$$

The aim of this quantum interference Bayesian Network is to simulate the averaged results reported in Table 1. More specifically, we are interested in simulating the value corresponding to the probability of the participant playing the second gamble, given that (s)he does not know the outcome of the first one. In Table 1, this value corresponds to 0.42 and cannot be obtained through classical probability theory, since the law of total probability is violated. In our model, we can tune the angle θ of the interference term in order to obtain such results. Calculations showed that in order to achieve a probability of 42%, $\cos(\theta)$ must be equal to -0.998853 , which corresponds to an angle of 177.3° or 3.09 radians.

$$G2_{00} = \frac{0.2950 + 2\sqrt{0.17}\sqrt{0.125}\cos(3.09)}{0.5 + 0.4915\cos(3.09)} = 0.42 \quad (56)$$

$$G2_{11} = \frac{0.2050 + 2\sqrt{0.08}\sqrt{0.125}\cos(3.09)}{0.5 + 0.4915\cos(3.09)} = 0.58 \quad (57)$$

This section described the application of the proposed quantum interference Bayesian Network to explain the puzzling findings in the two-stage gambling game, that could not be explained through classical probability theory. Equations 56 and 57 showed that the proposed approach is able to simulate the human decisions observed in the works of [71, 65, 50, 51].

6.3 The Impact of the Phase θ

The interference term that we show in this paper for the two-stage gambling game, was proposed by cognitive psychologists in order to explain their observations. That is, they manually tuned this parameter θ to fit their data. In this work, we look at this parameter from a different perspective: what happens to the quantum computed probabilities if we vary this angle θ ? Can better inferences be achieved? In this section, we make use of the probabilistic interference term proposed in the cognitive psychology literature [61, 22], and investigate the impact that the phase parameter θ can have, when computing quantum probabilities in the two step gambling game.

In the previous section, the general probability formula of a player willing to play the second gamble, given that the outcome of the first gamble was unknown, was given by Equation 54. In order to analyze the consequences of the angle θ in the final probability, we varied θ from 0 to 2π in steps of 0.0001 radians. The results obtained are discriminated in Figure 15.

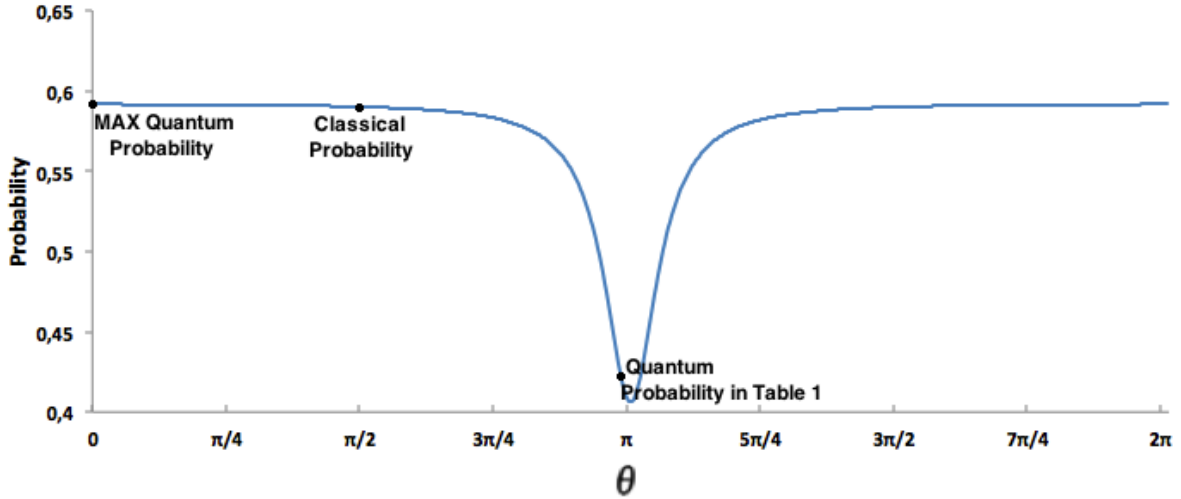


Figure 15: The various quantum probability values that can be achieved by varying the angle θ in Equation 54. Note that quantum probability can achieve much higher/lower values than the classical probability.

Figure 15 reveals that quantum probabilities can achieve much higher values than the classical probability theory. The quantum probability can reach a maximum of 0.5915 where the classical probability can only have a fixed value of 0.59.

Figure 15 is also supporting the quantum information processing theory already mentioned in Section 1 of this work: information is modeled via wave functions and therefore, they cannot be in a definite state (only when a final decision is made, a definite state emerges). One can look at all values that this parameter θ as all possible probabilities (or outcomes) that a player has when deciding to whether or not to play the second gamble. Since in quantum theory

we model the participant's beliefs by wave functions, then the superposed states can produce different waves coming from opposite directions that can crash into each other. When they crash, the waves can either unite or be destroyed. When they unite, it causes a constructive interference effect that will cause a bigger wave, leading to a maximum or minimum quantum probability value, depending on the phase of the wave (Figure 16). When the waves crash and are destroyed, then a destructive interference effect occurs (Figure 17).

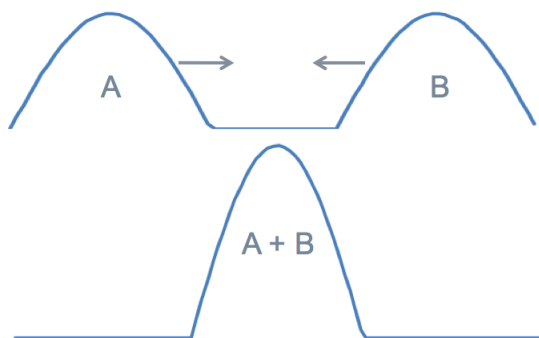


Figure 16: Example of constructive interference: two waves collide forming a bigger wave.

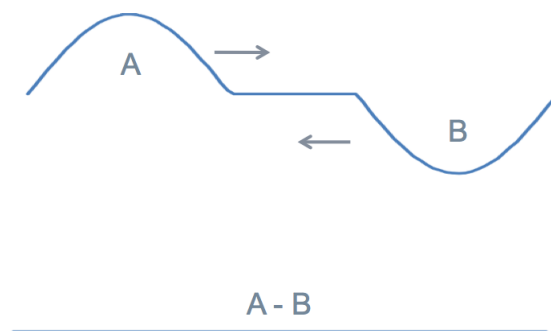


Figure 17: Example of destructive interference: two waves collide cancelling each other.

In conclusion, quantum probability enables the free choice of parameters in order to obtain a desired probability value. The two-stage gambling game is just a small example where the proposed model could be applied. In the next section, we turn to the task of burglary detection. We will analyse a more complex Bayesian Network that will determine the probability of a burglary occurring, given that the neighbours think that they heard an alarm.

7 Inference in More Complex Networks: The Burglar/Alarm Network

In this section, we compare classical inferences in Bayesian Networks with the proposed quantum model. The example that we will examine correspond to a modified version of the Burglar/Alarm network, which is very used in the scope of Artificial Intelligence [59, 63].

The network proposed in the book of [63] is inspired in the following scenario. Imagine that a person installed a new burglar alarm in his house. The alarm has a big credibility in what concerns detecting burglaries. The person has two neighbours: Mary and John. They have promised to call the person, whenever they think they heard the alarm. John always calls the police, when he hears the alarm, however, he sometimes confuses the sound of the alarm with the ringtone of his phone, resulting in a misleading phone call to the police. Mary, on the other hand, cannot hear the alarm very often, because she likes to hear very loud music. Given the evidence of who has or has not called the police, we want to represent in a Bayesian Network the probability of a burglary occurring [63].

Figure 18 represents a classical Bayesian Network to account for burglary detection. It's quantum counterpart corresponds to Figure 19. In order to make a fair comparison between both networks, the quantum Bayesian Network was built in the same way as the classical one,

Evidences	Pr(Alarm = t)	Pr(Burglar = t)	Pr(JohnCalls = t)	Pr(MaryCalls = t)
No Evidence	0.0347	0.0200	0.0795	0.0339
Alarm = t	1.0000	0.5479	0.9000	0.7000
Burglar = t	0.9500	1.0000	0.6655	0.8575
JohnCalls = t	0.3927	0.2158	1.0000	0.2810
MaryCalls = t	0.7155	0.3923	0.6582	1.0000
Alarm = t, Burglar = t	1.0000	1.0000	0.7000	0.9000
Alarm = t, JohnCalls = t	1.0000	0.5479	1.0000	0.7000
Alarm = t, MaryCalls = t	1.0000	0.5479	0.9000	1.0000
Burglar = t, JohnCalls = t	0.9971	1.0000	1.0000	0.6980
Burglar = t, MaryCalls = t	0.9992	1.0000	0.8994	1.0000
JohnCalls = t, MaryCalls = t	0.9784	0.5360	1.0000	1.0000

Table 3: Probabilities obtained when performing inference on the classical Bayesian Network of Figure 18.

but we replaced the real probability values by quantum complex amplitudes, just like proposed in [68] and [53].

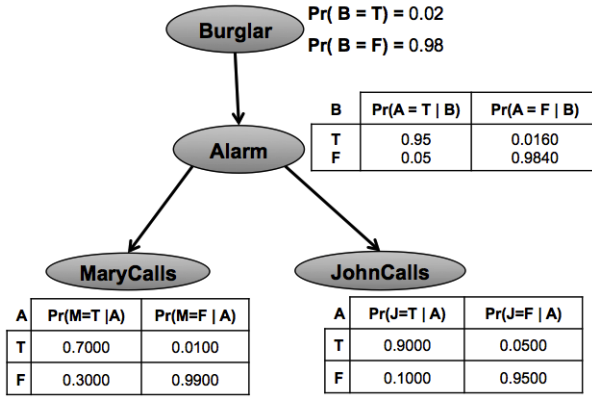
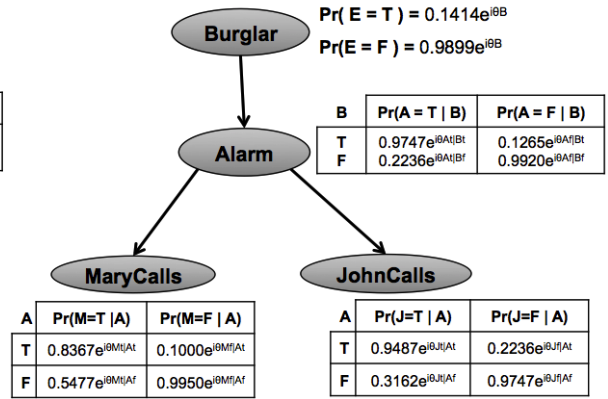


Figure 18: Burglar/Alarm classical Bayesian Network proposed in the book of [63]



We also performed a set of queries so we could compare the classical Bayesian Network with the proposed quantum interference Bayesian Network. Table 3 presents the final probabilities computed over the classical Bayesian Network of Figure 18 and Table 4 presents the results for the quantum Bayesian network (Figure 19).

In the experiment of the two-stage gambling game, we saw that the parameter θ plays an important role in determining the final probability value of a variable. For the Burglar/Alarm Bayesian Network, in order to find the best parameter for each query, we varied θ between 0 and 2π , in steps of 0.1 radians, and collected the θ that would maximize most variables in the network. Equation 58 represents the interference formula that we used to maximize the probability with respect to parameter θ .

$$Pr(A) = \operatorname{argmax}_{\theta} \alpha \left[\sum_{i=1}^N |\psi_i|^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N |\psi_i| |\psi_j| \cos(\theta_i - \theta_j) \right] \quad (58)$$

Evidences	Pr(Alarm = t)	Pr(Burglar = t)	Pr(JohnCalls = t)	Pr(MaryCalls = t)
No Evidence	0.1760	0.1179	0.1185	0.0889
Alarm = t	1.0000	0.5479	0.9000	0.7000
Burglar = t	0.9896	1.0000	0.9999	0.9791
JohnCalls = t	0.6596	0.8380	1.0000	0.8138
MaryCalls = t	0.9018	0.9998	0.9758	1.0000
Alarm = t, Burglar = t	1.0000	1.0000	0.9000	0.7000
Alarm = t, JohnCalls = t	1.0000	0.5479	1.0000	0.7000
Alarm = t, MaryCalls = t	1.0000	0.5479	0.9000	1.0000
Burglar = t, JohnCalls = t	0.9982	1.0000	1.0000	0.7390
Burglar = t, MaryCalls = t	0.9993	1.0000	0.9138	1.0000
JohnCalls = t, MaryCalls = t	0.9883	0.6632	1.0000	1.0000

Table 4: Probabilities obtained when performing inference on the quantum Bayesian Network of Figure 19.

In the next section, we will analyse the results specified in Tables 3 and 4 for each single query.

8 Discussion of Experimental Results

Tables 3 and 4 present the results obtained for different queries performed over the classical Bayesian Network (Figure 18) and the proposed quantum interference Bayesian Network (Figure 19), respectively.

Analysing the first row of both tables, when we provide no piece of evidence to the network, the proposed model was able to increase, on average, the probabilities of the query variables about 270.625%. In the case where nothing is observed, the proposed network achieves its maximum level of uncertainty: all variables are interfering with each other causing both destructive and constructive interference effects on the final probabilities. It is interesting to notice that when the classical probabilities computed for each query variable are very low, then the quantum Bayesian Network cannot greatly increase these probabilities. If the probabilities in a classical setting are very low, then the quantum Network is able to keep those probabilities low as well.

Figures 20-23 illustrate all possible values that each query variable could have, by varying the θ parameters. These graphs were plotted in the following way: for each query, we found the set of all θ 's that would lead to a maximum and a minimum probability value. We then compared the set of θ 's and realised that in the majority of the cases there were components that shared the same parameters. Given this situation, we fixed 6 of the parameters which were common and varied the remaining parameters, leading to a 3-dimensional graph.

The moment we start to provide pieces of evidence to the network, the uncertainty starts to decrease. Analyzing the situation where we observe that the *Alarm* variable is true, then an interesting phenomena occurs: the probabilities of the proposed quantum Bayesian Network collapse to the same probability values as in its classical counterpart. If one observes the variable *Alarm*, then the variables *Burglar*, *MaryCalls* and *JohnCalls* become independent of each other. This means that there is no possible way that these variables can interfere with each other. Since the variables do not provoke any interferences among them, then the interference term will be zero and will collapse to the classical probability. This independence phenomena

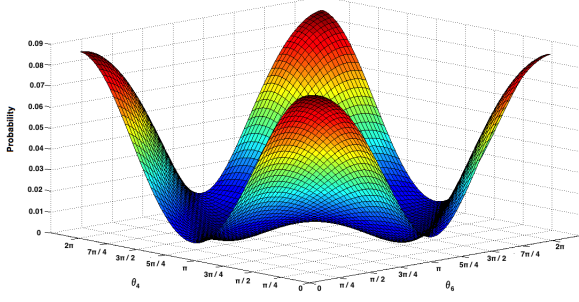


Figure 20: Possible probabilities when querying "MaryCalls = t" with no evidence. Parameters used were: $\{\theta_1, \theta_2, \theta_3, \theta_5, \theta_7, \theta_8\} \rightarrow \{0, 0, 0, 0, 3.1, 0\}$. Maximum probability for $\{\theta_1, \theta_2\} \rightarrow \{0, 3.1\}$.

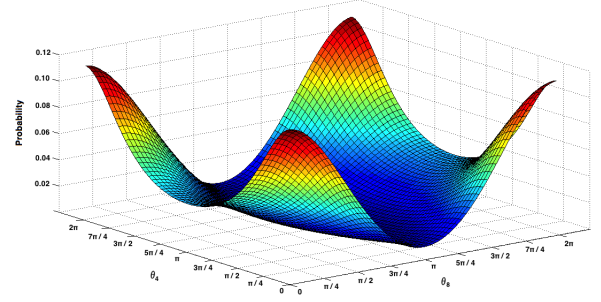


Figure 21: Possible probabilities when querying "Burglar = t" with no evidence. Parameters used were: $\{\theta_1, \theta_2, \theta_3, \theta_5, \theta_6, \theta_7\} \rightarrow \{0, 0, 0, 6.2, 0.1, 3.1\}$. Maximum probability for $\{\theta_4, \theta_8\} \rightarrow \{0, 3.2\}$.

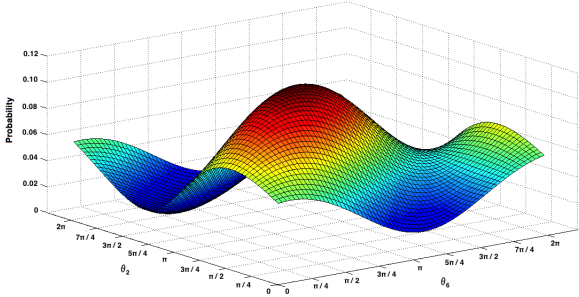


Figure 22: Possible probabilities when querying "JohnCalls = t" with no evidence. Parameters used were: $\{\theta_1, \theta_3, \theta_4, \theta_5, \theta_7, \theta_8\} \rightarrow \{1.9, 0, 2.3, 0.5, 4.5, 2.4\}$. Maximum probability for $\{\theta_2, \theta_6\} \rightarrow \{2.3, 5.5\}$.

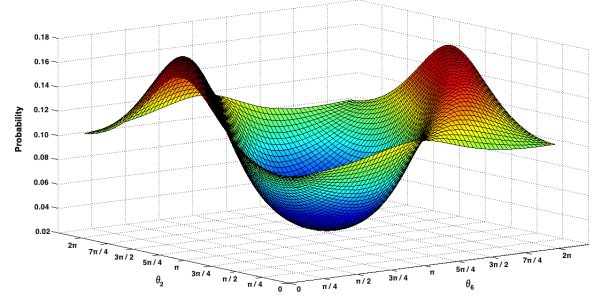


Figure 23: Possible probabilities when querying "Alarm = t" with no evidence. Parameters used were: $\{\theta_1, \theta_3, \theta_4, \theta_5, \theta_7, \theta_8\} \rightarrow \{0, 0, 0.8, 6.2, 3.1, 4.3\}$. Maximum probability for $\{\theta_2, \theta_6\} \rightarrow \{0.2, 0.5\}$.

in quantum Bayesian Networks has also been noticed in the work of [53]. For a mathematical proof of why this phenomena also occurs in a quantum setting, please refer to their work. This means that, whenever we observe the variable *Alarm*, the interference term will always be null and consequently all probabilities will be exactly the same as in a classical Bayesian Network inference.

In the case where we observe the *Burglar* variable, then, according to the scenario of the Bayesian Network, the *Alarm* variable should also increase. The variable *JohnCalls* is highly correlated with the variable *Alarm*. In its conditional probability table, there is a chance of 90% of the variable *JohnCalls* occurring when the variable *Alarm* is true. So in this situation, the quantum Bayesian Network is able to give more strength to that correlation. On average, the probabilities of all queries increased 22.8% when the variable *Burglar* is observed, when compared to the respective classical setting.

When we observe that the variable *JohnCalls* is true, then it is expected that the probability of the *Alarm* variable increases as well, according to the scenario of the Bayesian Network. When we observe the variable *MaryCalls* the variable *Alarm* increases even more when com-

pared to the situation where $JohnCalls = t$ or its classical counterpart. This means that the correlation between the *Alarm* variable and the variable *JohnCalls* is not as strong as when we observe *MaryCalls*. According to the scenario, John always calls the police, leading to many misleading calls. Consequently, the *Alarm* variable cannot be increased too much. However, when *MaryCalls* is observed, then, according to the Bayesian scenario, it is almost certain that she heard the alarm and therefore, a strict correlation exists between these two variables. The quantum Bayesian Network was able to represent the correlations between variables in a more realistic and reliable way than its classical counterpart.

Finally, when we start to provide 2 pieces of evidence to the network, then the uncertainty levels start to decrease. Consequently, the probabilities computed in a quantum setting start to converge to the ones computed in a classical Bayesian Network. This means that, there are two situations there the proposed quantum Bayesian Network converges to its classical counterpart: (1) when the variables of the network become independent of each other and (2) when there are very low levels of uncertainty, because too many evidences were provided to the network.

9 The Optimum Value for θ to Maximize Quantum Inferences in Bayesian Networks

In this section, we perform a study on the impact of the θ parameters in the context of medical decision making. Consider the Classical Bayesian Network of Figure 24, which corresponds to a slightly modified version of the Bayesian Network proposed in the book of [59]. Figure 25 corresponds to its quantum counterpart.

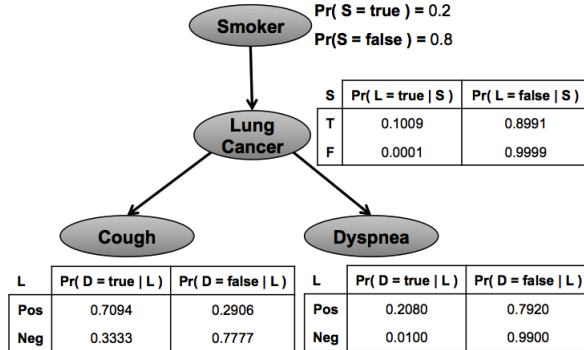


Figure 24: Classical representation of a Lung Cancer Bayesian Network inspired in the book of [59].

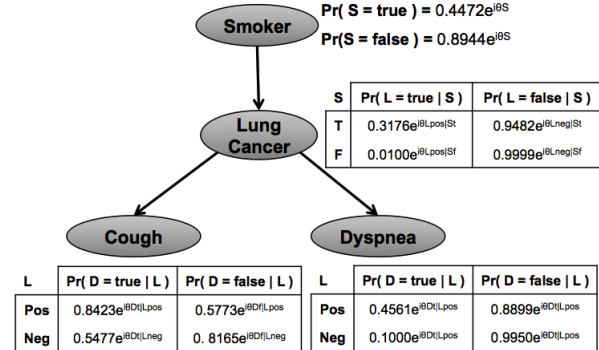


Figure 25: Quantum representation of a Lung Cancer Bayesian Network inspired in the book of [59].

Again, in order to determine the maximum probability value, we varied the θ parameters between 0 and 2π in steps of 0.1 radians. The method that we used to compute these parameters is described in Appendix A. We then performed the following queries: $\Pr(\text{Smoke} = \text{true})$, $\Pr(\text{Dyspnea} = \text{true})$, $\Pr(\text{Cough} = \text{high})$ and $\Pr(\text{Lung Cancer} = \text{positive})$. The results obtained are discriminated in Figure 26.

Analysing Figure 26, one can observe that again, the quantum probability values tend to overcome their classical counterparts. A special note goes to the quantum probabilities verified for the query $\Pr(\text{Cancer} = \text{positive})$. The increase verified in this probability was much higher

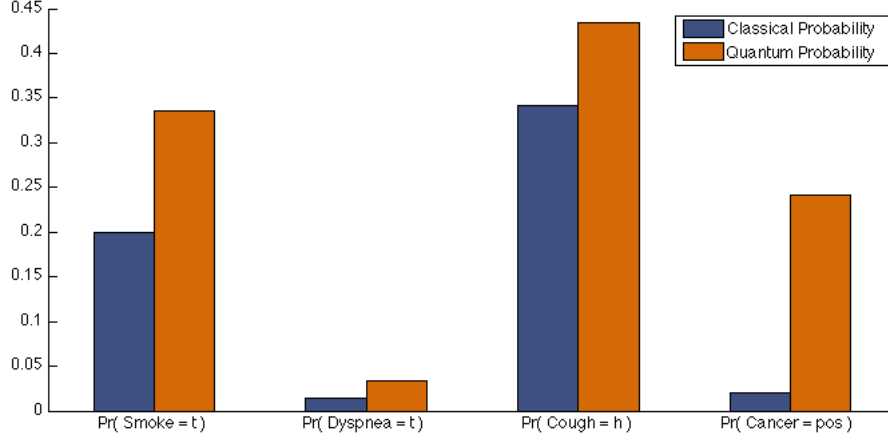


Figure 26: Classical and Quantum probabilities computed for different queries for the Lung Cancer Bayesian Network in Figure 24.

than any other value achieved by the other quantum probabilities. Since the *Lung_Cancer* random variable is located at the centered position of the Bayesian Network, then, when nothing is observed, its probability is influenced by the probabilities of all nodes in the network.

In order to determine the impact of the θ parameters, Table 5 shows the parameters that were used in the Burglar/Alarm and Lung Cancer Bayesian Networks.

Variables	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8
Alarm = t	0.00	0.20	0.00	0.80	6.20	0.50	3.10	4.30
Burglar = t	0.00	0.00	0.00	0.00	6.20	0.10	3.10	3.20
JohnCalls = t	1.90	2.30	0.00	2.30	0.50	5.50	4.50	2.40
MaryCalls = t	0.00	0.00	0.00	0.00	0.00	3.10	3.10	0.00
Cancer = t	3.20	3.20	3.20	3.20	3.20	3.20	0.10	0.00
Cough = t	0.00	0.00	0.00	0.00	3.10	3.20	0.10	0.00
Dyspnea = t	0.00	0.00	0.00	0.00	3.20	3.20	0.10	0.00
Smoke = t	0.00	0.00	0.00	0.00	3.20	3.20	0.10	0.00

Table 5: Optimum θ 's found for each variable from the burglar/alarm bayesian network (Figure 18) and from the lung cancer bayesian network (Figure 24).

Figures 27-30 also show all the possible probability values that the variables from the lung cancer Bayesian network can achieve, when nothing is known.

When we are at a maximum level of uncertainty, classical probabilities tend to assume that events are equiprobable, that is, the probability of their outcome is always the same no matter their context. Quantum theory, on the other hand, provides a more relaxed framework. When nothing is known, then there are more degrees of freedom that enable the outcome of any event to be any possible value. This way, by analysing the context of certain events (for example, a medical decision scenario, a gambling game, etc), the quantum parameters θ can be modelled in such a way that they can simulate reality more accurately and more precisely than its classical counterpart.

In this sense, in order to develop more accurate quantum Bayesian Networks, a study of the context of the problem is required in order to start the search for the optimum θ parameters.

Techniques to search for these optimum quantum parameters is still an open research question and an unexplored field in the literature of quantum cognition and quantum decision models.

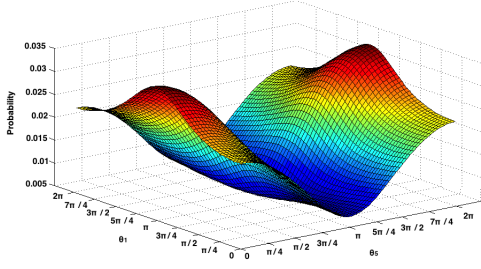


Figure 27: Possible probabilities when querying "Dyspnea = t" with no evidence.

Parameters used were:
 $\{\theta_2, \theta_3, \theta_4, \theta_6, \theta_7, \theta_8\} \rightarrow \{0, 0, 0, 3.2, 0.1, 0\}$. Maximum probability for $\{\theta_1, \theta_5\} \rightarrow \{0, 3.2\}$.

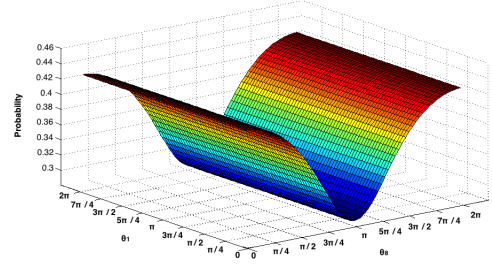


Figure 28: Possible probabilities when querying "Cough = t" with no evidence.

Parameters used were:
 $\{\theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7\} \rightarrow \{0, 0, 0, 3.1, 3.2, 0.1\}$. Maximum probability for $\{\theta_1, \theta_8\} \rightarrow \{0, 3.1\}$.

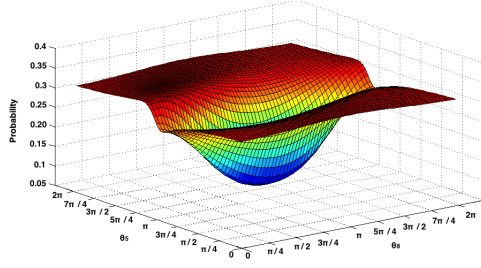


Figure 29: Possible probabilities when querying "Smoke = t" with no evidence.

Parameters used were:
 $\{\theta_1, \theta_2, \theta_3, \theta_4, \theta_6, \theta_7\} \rightarrow \{0, 0, 0, 0, 0.1, 0\}$.
 Maximum probability for $\{\theta_5, \theta_8\} \rightarrow \{0, 0.8\}$.

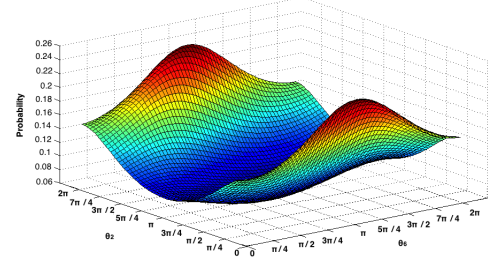


Figure 30: Possible probabilities when querying "Cancer = pos" with no evidence. Parameters used were:

$\{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_7\} \rightarrow \{3.2, 3.2, 3.2, 3.2, 3.2, 0.1\}$. Maximum probability for $\{\theta_2, \theta_6\} \rightarrow \{3.2, 3.2\}$.

10 Related Work

Since the preliminary findings of [69], cognitive scientists began to use alternative theories, such as quantum probability theory, in an attempt to explain the paradoxical findings that classical probability theory could not explain. Given the flexibility of the quantum probabilistic framework, several researchers applied these concepts in several fields outside of physics. In this section, we present some of the topics in which quantum probability has made some impacts in the literature.

10.1 Violations of Probability Theory

In what concerns violations of probability theory, there are many paradoxical situations where classical probability is unable to model. The most important paradoxes consist in violations in

the Sure Thing principle. Examples of such violations correspond to the Ellsberg paradox [32], the two-stage gambling game [71, 50, 51] and the prisoner's dilemma game [71, 26, 54, 18, 41, 25].

In what concerns the prisoner's dilemma game, many works in the literature have been proposed, which formalise the problem in a quantum approach. For instance, [61] proposed a quantum dynamic Markov Model for the prisoner's dilemma game. In their model, the authors represented the players beliefs and actions as a superposition in a quantum state vector. When the player knew the information about the opponent's action, the superposition state collapsed to a new state that was compatible with the action chosen by the opponent. When the player did not know about the opponent's move, then the system involved through the application of unitary transformations. In [15], the authors focus on irrational choices and developed a model based on the quantum superposition and interference principles for the prisoner dilemma game. After each play, the state is updated until it reaches an equilibrium. In [23], the authors analysed quantum conditional probabilities in the prisoner's dilemma game. There are also works that consider that the probabilities in the prisoner's dilemma game correspond to different contexts and these contexts can be incompatible. The authors argue that the probabilities observed in the prisoner's dilemma game were not classical, but were not quantum either in the sense of traditional quantum mechanics theory. So the author introduced the quantum-like approach and suggested a new trigonometric interference term [47, 48]. [72] proposed quantum model which incorporated entangled decisions and explained the violations of the Sure Thing principle through interference effects between intentions. In [1], the authors explored the impact of these violations in economics and analyse the two-stage gambling game and the prisoner's dilemma game under a quantum probabilistic point of view, by proposing a quantum Markov Model to explain the paradoxical observations in these games. Finally, [14] proposed a quantum bayesian updating scheme based on the formalisms of quantum mechanics that represents mental states on a Hilbert state. Through the usage of projections, they were able to introduce a model that could explain the paradoxical findings of the two-stage gambling game [71]. Another similar work that compares Bayes rule to its quantum counterpart corresponds to [20].

10.2 Conjunction and Diskunction Errors

Other situations, in which the laws of probability are being violated, correspond to disjunction or conjunction fallacies. A conjunction error occurs when it is assumed that specific conditions are more probable than a single general one. A disjunction error, on the other hand, consists in assuming that the disjunction of two events is at least as likely as either of the events individually. Several works in the literature used quantum probabilistic models to address these fallacies. For instance, [34] developed a quantum probabilistic model to explain fallacies when making preferences over a set of events, more specifically, the conjunction fallacy. [19] also focused on quantum models to explain conjunction and [26] focused on disjunction fallacies. The first experiments where these fallacies were observed to occur were performed by [71]. An alternative model for conjunction and disjunction errors correspond to [70]. The authors used support theory as a subjective probabilistic model that describes the way people make probability judgments. The main problem of their model is that support theory is able to successfully describe conjunction errors, but fails at explaining disjunction errors. Other models correspond to [37, 36]. [72] also addressed the problem of conjunction and disjunction fallacies. They proposed quantum model which incorporated entangled decisions. [44] also modelled mental

processes through quantum probabilities, where the interference process plays an important role in the process of recognising images[24]. Other interesting works of this author applying similar quantum formalisms correspond to [47, 46, 45, 67].

10.3 Quantum Probabilistic Graphical Models

In what concerns quantum probabilistic graphical models, there are various contributions. For example, in [68] the model proposed by the author is exactly the same as a classical Bayesian Network. The only difference is that it uses complex numbers in the conditional probability tables instead of real values. A similar model has been proposed by [57], but for Markov Networks. In [53], the author proposed a quantum Bayesian Network by replacing the classical formulas used to perform the inference process by their quantum counterpart. [22], [18] and [21] proposed a quantum dynamic Markov model based on the findings of cognitive psychologists and interference terms.

10.4 General Applications of Quantum Probabilities

One of the first and most influential models that applied quantum probability for decision making belongs to [3]. In their work, the authors proposed the ε -model, which can be defined as a quantum machine that corresponds to a two dimensional Hilbert space. Given an event in this quantum machine one can compute quantum like probabilities. The model also makes use of a parameter ε that measures the total amount of uncertainty when computing the probabilities. Other works related to this model can be found in [8, 9, 7, 2, 10, 11, 2, 12, 13, 35, 4].

[14] proposed a quantum bayesian updating scheme based on the formalisms of quantum mechanics that represents mental states on a Hilbert state. Through the usage of projections, they were able to introduce a model that could explain the paradoxical findings of the two-stage gambling game [71]. Another similar work that compares Bayes rule to its quantum counterpart corresponds to [20].

There is also an increasing interest in applying quantum like models to subjects outside of psychology. In game theory there have been works exploring the uncertainties of a player towards another through quantum probabilistic models. For instance, [15] focus on irrational choices and developed a model based on the quantum superposition and interference principles for the Prisoner Dilemma game. After each play, the state is updated until it reaches an equilibrium. [23] used the data collected by [71] and analysed quantum conditional probabilities. [56] generalised the classical expected utility and proposed a quantum projective expected utility function that does not violate the laws of classical probability theory and can accommodate Allais and Ellsberg paradoxes. [27] made a formalisation of the structure of quantum probability theory and [28] also applied these formalists to expected utilities. [60] applied quantum theory to the fields of economics and game theory and developed a principle to minimise financial risks [6] modelled how a person updates its beliefs in the liar's paradox. [44] also modelled mental processes through quantum probabilities, where the interference process plays an important role in the process of recognising images[24]. Other interesting works of this author applying similar quantum formalisms correspond to [47, 48, 46, 45, 42, 39]. [52] developed a quantum mechanics based framework in order to model agents preferences under uncertainty.

More recently, [62] published a work that generated discussions over the scientific literature, about whether or not, quantum probability can provide a new direction to compute quan-

tum probabilistic inferences. In their work, they summarised the main areas where quantum probability has made some impact (conjunction/disjunction errors, sure thing principle, etc). A deep analysis of the differences between classical and quantum probability is also made in this work.

[5] also shares a similar opinion that quantum probability can, in fact, provide a new framework to explain several situations in which the laws of probability are violated. However, they defend that the usage of quantum probability should be more close to its real meaning in quantum mechanics and proposes to explain the phenomena, such as contextually, entanglement, observables and Flock spaces.

The work of, [62] has stimulated some discussions on the implications of quantum probability in functional brain networks. For instance, [66] have discussed the similarities between the vector representations of quantum states with a vector symbolic architecture, which is used to model realistic biological neural models. In this sense, biological neural models can also incorporate the quantum probabilistic framework and also introduce quantum interference effects, as a mathematical and geometric alternative framework, without taking into account the true meaning that is given these effects under quantum physics. Other authors also support this idea of representing neuronal model using the quantum probabilistic framework [16, 38, 30, 29].

11 Conclusions

This work was motivated by the preliminary experiments of [71] about violations of the classical probability theory on the sure thing principle. This principle states that if one chooses action A over B in a state of the world X , and if one also chooses action A over B under the complementary state of the world X , then one should always choose action A over B , even when the state of the world is unspecified. When humans need to make decisions under risk, several heuristics are used, since humans cannot process large amounts of data. These decisions coupled with heuristics lead to violations on the law of total probability.

Recent work in cognitive psychology revealed that quantum probability theory provides another method of computing probabilities without falling into the restrictions that classical probability have in modelling cognitive systems of decision making. Quantum probability theory can also be seen as a generalisation of classical probability theory, since it also includes the classical probabilities as a special case (when the interference term is zero).

The main difference between quantum and classical probability lies in the fact that on quantum probability we are constantly updating some beliefs when making a decision, while in classical probability all beliefs are assumed to have a definite value before a decision is made, and this value is the outcome of the decision [3].

The main research question for this work was how could these quantum probabilities affect probabilistic graphical models, such as Bayesian Networks, since many of nowadays decision making systems are based on such structures (medical diagnosis, spam filtering, image segmentation, etc).

In this work, we proposed a novel Bayesian Network for the Computer Science community based on quantum probabilities. Our method can accommodate puzzling observations that the classical probability failed to explain (for instance, the two-step gambling game). When the nodes of the proposed Bayesian Network are represented as a superposition state, then one can look at this state as many waves moving across in different directions. These waves can crash

into each other causing waves to be bigger or to cancel each other. This is the interference phenomena that the proposed Bayesian Network offers and that has direct implications when making inferences. Therefore, the proposed network represents and simulates quantum aspects motivated by Feynman's path integrals.

Experimental results revealed that the proposed quantum Bayesian Network enables many degrees of freedom in choosing the final outcome of the probabilities. If we had a real scenario, with real observations, one could use the present model to fit it to the observed data, by simply tuning the parameter θ . This parameter can open a door into machine learning approaches. Learning algorithms using the proposed method might produce better prediction models, since the quantum probability amplitudes are able to fully represent real word data. For future work, we intend to explore machine learning algorithms under quantum probabilistic graphical model formalisms.

The overall results also suggested that when the classical probability of some variable is already high, then the quantum probability tends to increase it even more. When the classical probability is very low, then the proposed model tends to lower it.

When there are many unobserved nodes in the network then the levels of uncertainty are very high. But, in the opposite scenario, when there are very few unobserved nodes, then the proposed quantum model tends to collapse into its classical counterpart, since the uncertainty levels are very low.

The proposed Bayesian Network can integrate human thoughts by representing a person's beliefs in an N-dimensional unit length quantum state vector. In the same way, the proposed quantum structure is general enough to represent any other context in which there is a need to formalise uncertainty, including prediction problems in data fusion. In the context of Bayesian Networks, data fusion is introduced in the work of [58]. The author argues that, just like people, Bayesian Networks are structures that integrate data from multiple sources of evidence and enable the generation of a coherent interpretation of that data through a reasoning process. The fusion of all these multiple data sources can be done using Bayes theorem. When a data source is unknown, then the Bayes rule is extended in order to sum out all possible values of the probability distribution representing the unknown data source. The proposed Quantum Bayesian Network takes advantage of these uncertainties by representing them in a superposition state, which enables the fusion of the data sources through quantum interference effects. These effects produce changes in the final likelihoods of the outcomes and provide a promising way to perform predictions more accurately according to reality. So, the Quantum Bayesian Network that is proposed in this work is potentially relevant and applicable in any behavioural situation in which uncertainty is involved.

References

- [1] Luigi Accardi, Andrei Khrennikov, and Masanori Ohya. Quantum markov model for data from shafir-tversky experiments in cognitive psychology. *Journal of Open Systems and Information Dynamics*, 14:371–385, 2009.
- [2] Diederik Aerts. Quantum structures: An attempt to explain the origin of their appearance in nature. *International Journal of Theoretical Physics*, 34:1–22, 1995.

- [3] Diederik Aerts. Quantum structure in cognition. *Journal of Mathematical Psychology*, 53:314–348, 2009.
- [4] Diederik Aerts, Jan Broekaert, and Liane Gabora. A case for applying an abstracted quantum formalism to cognition. *Journal of New Ideas in Psychology*, 29:136–146, 2011.
- [5] Diederik Aerts, Jan Broekaert, Liane Gabora, and Sandro Sozzo. Quantum structure and human thought. *Journal of Brain and Behavioral Sciences*, 36:274–276, 2013.
- [6] Diederik Aerts, Jan Broekaert, and Sonja Smets. A quantum structure description of the liar paradox. *International Journal of Theoretical Physics*, 38:3231–3239, 2004.
- [7] Diederik Aerts and Thomas Durt. Quantum, classical and intermediate, an illustrative example. *Foundations of Physics*, 24:1353–1369, 1994.
- [8] Diederik Aerts, Thomas Durt, and Bruno Van Bogaert. A physical example of quantum fuzzy sets and the classical limit. *Tatra Mountains Mathematical Publications*, 1:5–15, 1993.
- [9] Diederik Aerts, Thomas Durt, and Bruno Van Bogaert. Quantum probability, the classical limit and nonlocality. In *Proceedings of the International Symposium on the Foundations of Modern Physics*, 1993.
- [10] Diederik Aerts and Liane Gabora. A theory of concepts and their combinations i: The structure of the sets of contexts and properties. *Journal of Kybernetes*, 34:167–191, 2005.
- [11] Diederik Aerts and Liane Gabora. A theory of concepts and their combinations ii: A hilbert space representation. *Journal of Kybernetes*, 34:192–221, 2005.
- [12] Sven Aerts. Conditinal probabilities with a quantal and kolmogorovian limit. *International Journal of Theoretical Physics*, 35:2245, 1996.
- [13] Sven Aerts. Interactive probability models: Inverse problems on the sphere. *International Journal of Theoretical Physics*, 37:305–309, 1998.
- [14] Masanari Asano, Irina Basieva, Andrei Khrennikov, Masanori Ohya, and Yoshiharu Tanaka. Quantum-like generalization of the bayesian updating scheme for objective and subjective mental uncertainties. *Journal of Mathematical Psychology*, 56:166–175, 2012.
- [15] Masanari Asano, Masanori Ohya, and Andrei Khrennikov. Quantum-like model for decision making process in two players game, a non-kolmogorovian model. *Journal of Foundations of Physics*, 41:538–548, 2010.
- [16] Arpan Banerjee and Barry Horwitz. Can quantum probability help analyze the behaviour of functional brain networks? *Journal of Behavioral and*, 36:278–279, 2013.
- [17] Jerome Busemeyer and Peter Bruza. *Quantum Model of Cognition and Decision*. Cambridge University Press, 2012.
- [18] Jerome Busemeyer, Marvin Matthew, and Zheng Wang. A quantum information processing explanation of disjunction effects. In *Proceedings of the 28th Annual COnference of the Cognitive Science Society*, 2006.

- [19] Jerome Busemeyer, Emmanuel Pothos, Riccardo Franco, and Jennifer Trueblood. A quantum theoretical explanation for probability judgment errors. *Journal of Psychology Review*, 118:193–218, 2011.
- [20] Jerome Busemeyer and Jennifer Trueblood. Comparison of quantum and bayesian inference models. In *Proceedings of the 3rd International Symposium on QUantum Interaction*, 2009.
- [21] Jerome Busemeyer, Zheng Wang, and Ariane Lambert-Mogiliansky. Empirical comparison of markov and quantum models of decision making. *Journal of Mathematical Psychology*, 53:423–433, 2009.
- [22] Jerome Busemeyer, Zheng Wang, and James Townsend. Quantum dynamics of human decision making. *Journal of Mathematical Psychology*, 50:220–241, 2006.
- [23] Taksu Cheon and Taiki Tahahashi. Interference and inequality in quantum decision theory. *Journal of Physics Letters A*, 375:100–104, 2010.
- [24] Elio Conte, Andrei Khrennikov, Orlando Todarello, Antonio Federici, Leonardo Mendolicchio, and Joseph Zbilut. Mental states follow quantum mechanics during perception and cognition of ambiguous figures. *Journal of Open Systems and Information Dynamics*, 16:1–17, 2009.
- [25] Elio Conte, Orlando Todarello, Antonio Federici, Francesco Vitiello, Michele Lopane, Andrei Khrennikov, and Joseph Zbilut. Some remarks on an experiment suggesting quantum like behavior of cognitive entities and formulation of an abstract quantum mechanical formalism to describe cognitive entity and its dynamics. *Journal of Chaos, Solitons and Fractals*, 31:1076–1088, 2007.
- [26] Ratchel Croson. The disjunction effect and reason-based choice in games. *Journal of Organizational and Human Decision Processes*, 80:118–133, 1999.
- [27] Vladimir Danilov and Ariane Lambert-Mogiliansky. Measurable systems and behavioral sciences. *Journal of Mathematical Social Sciences*, 55:315–340, 2008.
- [28] Vladimir Danilov and Ariane Lambert-Mogiliansky. Expected utility theory under non-classical uncertainty. *Journal of Theory and Decision*, 68:25–47, 2010.
- [29] José Acacio de Barros. Quantum-like model of behavioral response computation using neural oscillators. *Journal of BioSystems*, 110:171–182, 2012.
- [30] Jose Acacio de Barros and Patrick Suppes. Quantum mechanics, interference, and the brain. *Journal of Mathematical Psychology*, 53:306–313, 2009.
- [31] Morris DeGroot and Mark Schervish. *Probability and Statistics*. Pearson Education (4th Edition), 2011.
- [32] Daniel Ellsberg. Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics*, 75:643–669, 1961.

- [33] Richard Feynman, Robert Leighton, and Mathew Sands. *The Feynman Lectures on Physics: Quantum Mechanics*. Addison-Wesley, 1965.
- [34] Riccardo Franco. The conjunction fallacy and interference effects. *Journal of Mathematical Psychology*, 53:415–422, 2009.
- [35] Liane Gabora and Diederik Aerts. Contextualizing concepts using a mathematical generalization of the quantum formalism. *Journal of Experimental & Theoretical Artificial Intelligence*, 14:327–358, 2002.
- [36] Gerd Gigerenzer and Daniel G. Goldstein. Reasoning the fast and frugal way: Models of bounded rationality. *Journal of Psychological Review*, 103:650–669, 1996.
- [37] Gerd Gigerenzer and Ulrich Hoffrage. How to improve bayesian reasoning without instruction: Frequency formats. *Journal of Psychological Review*, 102:684–704, 1995.
- [38] Stuart Hameroff. How quantum brain biology can rescue conscious free will. *Journal of Frontiers in Integrative Neuroscience*, 6:1–17, 2012.
- [39] Emmanuel Haven and Andrei Khrennikov. *Quantum Social Science*. Cambridge University Press, 2013.
- [40] Mika Hirvensalo. *Quantum Computing (Second Edition)*. Springer, 2003.
- [41] Evgenia Hristova and Maurice Grinberg. Disjunction effect in prisoner’s dilemma: Evidences from an eye-tracking study. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, 2008.
- [42] Andrei Khrennikov. Classical and quantum mechanics on information spaces with applications to cognitive psychological, social and anomalous phenomena. *Foundations of Physics*, 29:1065–1098, 1999.
- [43] Andrei Khrennikov. Interference of probabilities in the classical probabilistic framework. *Journal of Fuzzy Sets and Systems*, 155:4–17, 2005.
- [44] Andrei Khrennikov. Quantum-like brain: Interference of minds. *Journal of BioSystems*, 84:225–241, 2006.
- [45] Andrei Khrennikov. Can quantum information be processed by macroscopic systems? *Quantum Information Processing*, 6:401–429, 2007.
- [46] Andrei Khrennikov. Description of composite quantum systems by means of classical random fields. *Foundations of Physics*, 40:1051–1064, 2009.
- [47] Andrei Khrennikov. Quantum-like model of cognitive decision making and information processing. *Journal of BioSystems*, 95:179–187, 2009.
- [48] Andrei Khrennikov and Emmanuel Haven. Quantum mechanics and violations of the sure-thing principle: The use of probability interference and other concepts. *Journal of Mathematical Psychology*, 53:378–388, 2009.

- [49] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, 2009.
- [50] Anton Kuhberger, Dagmara Komunska, and Perner Josef. The disjunction effect: Does it exist for two-step gambles? *Organizational Behavior and Human Decision Processes*, 85:250–264, 2001.
- [51] Charles Lambdin and Charles Burdsal. The disjunction effect reexamined: Relevant methodological issues and the fallacy of unspecified percentage comparisons. *Organizational Behavior and Human Decision Processes*, 103:268–276, 2007.
- [52] Ariane Lambert-Mogiliansky, Shmuel Zamir, and Herve Zwirn. Type indeterminacy: A model for the kt(kahneman-tversky)-man. *Journal of Mathematical Psychology*, 53:349–361, 2009.
- [53] Mathew Leifer and David Poulin. Quantum graphical models and belief propagation. *Annals of Physics Journal*, 323:1899–1946, 2008. eprint arxiv-ph/0708.1337 [quant-ph].
- [54] Shu Li and John Taplin. Examining whether there is a disjunction effect in prisoner’s dilemma game. *Chinese Journal of Psychology*, 44:25–46, 2002.
- [55] Donald Metzler and W. Bruce Croft. Linear feature-based models for information retrieval. *Information Retrieval*, 16(28):1–23, 2007.
- [56] Pierfrancesco La Mura. Projective expected utility. *Journal of Mathematical Psychology*, 53:408–414, 2009.
- [57] Pierfrancesco La Mura and Lukasz Swiatczak. Markov entangled networks. Technical report, AAAI Spring Symposium, 2007.
- [58] Judea Pearl. Fusion, propagation, and structuring in belief networks. *Journal of Artificial Intelligence*, 29:241–288, 1986.
- [59] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, 1988.
- [60] Edward Piotrowski. Quantum-like approach to financial risk: Quantum anthropic principle. *Acta Physica Polonica B*, 32:3873–3879, 2001.
- [61] Emmanuel Pothos and Jerome Busemeyer. A quantum probability explanation for violations of rational decision theory. *Proceedings of the Royal Society B*, 276:2171–2178, 2009.
- [62] Emmanuel Pothos and Jerome Busemeyer. Can quantum probability provide a new direction for cognitive modeling? *Journal of Brain and Behavioral Science*, 36:255–327, 2013.
- [63] Stuart Russel and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson Education (3rd Edition), 2010.
- [64] Leonard Savage. *The Foundations of Statistics*. John Wiley, 1954.

- [65] Eldar Shafir and Amos Tversky. Thinking through uncertainty: nonconsequential reasoning and choice. *Cognitive Psychology*, 24:449–474, 1992.
- [66] Terrence Stewart and Chris Eliasmith. Realistic neurons can compute the operations needed by quantum probability theory and other vector symbolic architectures. *Journal of Brain and Behavioral Science*, 36:307–308, 2013.
- [67] Katya Tentori, Vincenzo Crupi, and Selena Russo. On the determinants of the conjunction fallacy: Confirmation versus probability. *Journal of Experimental Psychology: General*, 142:23555, 2013.
- [68] Robert Tucci. Quantum bayesian nets. *International Journal of Modern Physics B*, 9:295–337, 1995.
- [69] Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185:1124–1131, 1974.
- [70] Amos Tversky and Derek Koehler. Support theory: A nonextensional representation of subjective probability. *Journal of Psychological Review*, 10:547–567, 1994.
- [71] Amos Tversky and Eldar Shafir. The disjunction effect in choice under uncertainty. *Journal of Psychological Science*, 3:305–309, 1992.
- [72] Vyacheslav Yukalov and Didier Sornette. Decision theory with prospect interference and entanglement. *Theory and Decision*, 70:283–328, 2011.

A A Grid Search Approach to Find Quantum Parameters

The most naive approach of parameter search is the grid search method. In this approach, it is placed a grid over the parameter space and the data is evaluated at every grid intersection, returning the parameters, which lead to the maximum performance of an algorithm [55]. However, grid search has the problem of being unbounded, since an infinite set of parameters are available to be tested. In the scope of this work, this is not a problem, because we always have a boundary. The values of a cosine function can all be specified for angles between the range $[0, 2\pi]$.

The grid search approach is a very naive method for finding parameters in the parameter space. In fact, there are several advanced parameter search algorithms in the literature, which do not have a heavy computational cost. The motivation behind the usage of the grid search approach in this work is that the computational time required by any of the other advanced methods is almost the same as using grid search. In addition, many of the advanced search parameter methods in the literature perform approximations, which can be avoided by the direct parameter search of the grid search approach [55].

Algorithm 1 Grid Search to find θ parameters

Require: classical probability when it is true ($\Pr(A = t)$), Pr_{Ct} ,
interference term correspondent to Pr_{Ct} , $Interference_t$,
classical probability when it is true ($\Pr(A = f)$), Pr_{Cf} ,
interference term correspondent to Pr_{Cf} , $Interference_f$,

Ensure: Maximum quantum probabilities,
List of θ parameters that maximise the quantum probability

```
1:  $max\_probability \leftarrow -1$       // initialise variable that will store the maximum probability
   found

2:  $list\_parameters \leftarrow empty$  // initialise variable that will store all  $\theta$  parameters that
   maximise the probability

3: // Create a for-cycle for each quantum parameter necessary to compute the probability
4: for  $\theta_1 = 0; \theta_1 \leq 2\pi; \theta_1 = \theta_1 + 0.01$  do
5:   for  $\theta_2 = 0; \theta_2 \leq 2\pi; \theta_2 = \theta_2 + 0.01$  do
6:     ... // Add other for-cycles to compute the other  $\theta$  parameters

7:      $Pr\_positive \leftarrow Pr_{Ct} + Interference_t \cdot \cos(\theta_1 - \theta_2);$ 
8:      $Pr\_negative \leftarrow Pr_{Cf} + Interference_f \cdot \cos(\theta_1 - \theta_2);$ 

9:      $Pr\_positive\_norm \leftarrow Pr\_positive / (Pr\_positive + Pr\_negative);$ 
10:     $Pr\_negative\_norm \leftarrow Pr\_negative / (Pr\_positive + Pr\_negative);$ 

11:    if  $Pr\_positive\_norm > max\_probability$  then
12:       $max\_probability \leftarrow Pr\_positive\_norm;$ 
13:       $list\_parameters \leftarrow [\theta_1, \theta_2];$ 
14:    end if

15:  end for
16: end for

17: return  $max\_probability, list\_parameters;$ 
```
